

Smart Innovation, Systems and Technologies 279

José Luís Reis
Eduardo Parra López
Luiz Moutinho
José Paulo Marques dos Santos *Editors*



Marketing and Smart Technologies

Proceedings of ICMarTech 2021,
Volume 1



 Springer

Smart Innovation, Systems and Technologies

Volume 279

Series Editors

Robert J. Howlett, Bournemouth University and KES International,
Shoreham-by-Sea, UK

Lakhmi C. Jain, KES International, Shoreham-by-Sea, UK

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at <https://link.springer.com/bookseries/8767>

José Luís Reis · Eduardo Parra López ·
Luiz Moutinho · José Paulo Marques dos Santos
Editors

Marketing and Smart Technologies

Proceedings of ICMarTech 2021, Volume 1

Editors

José Luís Reis
University of Maia - ISMAI
Maia, Portugal

Eduardo Parra López
University of La Laguna
Tenerife, Spain

Luiz Moutinho
Suffolk Business School
University of Suffolk
Ipswich, UK

José Paulo Marques dos Santos
University of Maia - ISMAI
Maia, Portugal

ISSN 2190-3018

ISSN 2190-3026 (electronic)

Smart Innovation, Systems and Technologies

ISBN 978-981-16-9267-3

ISBN 978-981-16-9268-0 (eBook)

<https://doi.org/10.1007/978-981-16-9268-0>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022, corrected publication 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

This book is composed by the papers accepted for presentation and discussion at The 2021 International Conference on Marketing and Technologies (ICMarkTech'21). This conference had the support of the University of La Laguna. It took place at Tenerife, Spain, December 2–4, 2021.

The 2021 International Conference on Marketing and Technologies (ICMarkTech'21) is an international forum for researchers and professionals to present and discuss the latest innovations, trends, results, experiences and concerns in the various fields of marketing and technologies related to it.

The Program Committee of ICMarkTech'21 was composed of a multidisciplinary group of 285 experts and those who are intimately concerned with marketing and technologies. They have had the responsibility for evaluating, in a 'double-blind review' process, the papers received for each of the main themes proposed for the conference: (A) Artificial Intelligence Applied in Marketing; (B) Virtual and Augmented Reality in Marketing; (C) Business Intelligence Databases and Marketing; (D) Data Mining and Big Data—Marketing Data Science; (E) Web Marketing, e-commerce and v-commerce; (F) Social Media and Networking; (G) Omnichannel and Marketing Communication; (H) Marketing, Geomarketing and IOT; (I) Marketing Automation and Marketing Inbound; (J) Machine Learning Applied to Marketing; (K) Customer Data Management and CRM; (L) Neuro-marketing Technologies; (M) Mobile Marketing and Wearable Technologies; (N) Gamification Technologies to Marketing; (O) Blockchain Applied to Marketing; (P) Technologies Applied to Tourism Marketing; (Q) Digital Marketing and Branding; and (R) Innovative Business Models and Applications for Smart Cities.

ICMarkTech'21 received about 200 contributions from 30 countries around the world. The papers accepted for presentation and discussion at the conference are published by Springer (this book, Volume 1 and Volume 2) and will be submitted for indexing by ISI, EI-Compendex, Scopus, DBLP and/or Google Scholar, among others.

We acknowledge all of those that contributed to the staging of ICMarkTech'21 (authors, committees, workshop organizers and sponsors). We deeply appreciate their involvement and support that were crucial for the success of ICMarkTech'21.

Maia, Portugal
Tenerife, Spain
Ipswich, UK
Maia, Portugal
December 2021

José Luís Reis
Eduardo Parra López
Luiz Moutinho
José Paulo Marques dos Santos

Contents

Artificial Intelligence Applied in Marketing	
Forecasting Hotel-booking Cancellations Using Personal Name Records: An Artificial Intelligence Approach	3
Eleazar C.-Sánchez, Agustín J. Sánchez-Medina, and Laura Romero-Domínguez	
Managerial Decisions in Marketing: The Individual Perception of Explainable Artificial Intelligence	15
Gioia V. Volkmar	
Find Me if You Can! Identification of Services on Websites by Human Beings and Artificial Intelligence	23
Christian Stadlmann, Gábor Berend, Alexandra Fratrič, Zsolt Szántó, David Aufreiter, Margarethe Ueberwimmer, and Stefan Mang	
Under Which Conditions Are Humans Motivated to Delegate Tasks to AI? A Taxonomy on the Human Emotional State Driving the Motivation for AI Delegation	37
Anna Bouwer	
Roadmapping Collaborative Exploitation and Marketing of an AI-Based Knowledge Platform	55
Andrzej M. J. Skulimowski	
Business Intelligence Databases and Marketing	
Does Concept Recall in Brand Image Show High Loyalty? An Experimental Study on the Apple MacBook	69
Takumi Kato	
What Makes a Movie Get Success? A Visual Analytics Approach	79
Bruno Vaz, Maria de Fátima Barros, Maria João Lavoura, and Álvaro Figueira	

Impact of COVID-19 and Rapid Response of Small Restaurants in Thailand	93
Saruda Sunthornpan and Sadayo Hirata	
Improving Communication with Media: Portuguese National Public Police Case	105
Sónia M. A. Morgado	
Database Encryption Performance Impact on PostgreSQL and MongoDB	121
Marcelo Costa, Margarida Rodrigues, Pedro Baptista, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi	
Multi-DB Data Streaming on Polyglot Systems	129
Luis Oliveira, João Brito, Frederico Cá, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi	
OSSpal Qualitative and Quantitative Comparison: Couchbase, CouchDB, and MongoDB	141
Tiago Cardoso, João Penela, André Rosa, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi	
Security Analysis in the Architecture of the ATM Service	151
Gissela Quiña, Daisy Imbaquingo Esparza, Tatyana Saltos-Echeverria, José Jácome León, and MacArthur Ortega	
Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution	171
Filomena Castro Lopes, Paula Morais, and Abílio Cardoso	
An Application of Open Data in Public Administrations: The Lanzarote Tourism Indicator Dashboard	181
Christian González-Martel and José Manuel Cazorla-Artiles	
Structuring Best Practices of Search Engine Optimization for Webpages	191
Riaze Miguel Ribeiro Issá and Jose Paulo Marques dos Santos	
Data Mining and Big Data—Marketing Data Science	
The Influence of Emotions and Score Attributes on Online Ratings in P2P Accommodation: A Data Mining Study	213
E. Bigné, M. L. Fuentes-Medina, and Sandra Morini-Marrero	
Sensing the Impact of COVID-19 Restrictions from Online Reviews: The Cases of London and Paris Unveiled Through Text Mining	223
Bruno Silva, Sérgio Moro, and Catarina Marques	

Sources Used to Data-Driven Decision-Making in Tourism Management. Identifying the Main Areas of Research 233
 Juan Vidal, Ramón A. Carrasco, María F. Blasco, and Manuel J. Cobo

Assessing Corporate Reputation from Online Employee Reviews 243
 R. E. Loke and IJ. A. A. Steentjes

Online Reviews of Discount Products: The Case of Steam 259
 Bruno Sista, Beatriz Casais, and Nuno Moutinho

Cough Sound Identification: An Approach Based on Ensemble Learning 269
 Christian Salamea-Palacios, Javier Guaña-Moya, Tarquino Sanchez, Xavier Calderón, and David Naranjo

Advantages and Benefits of Big Data in Business Communication 279
 Mario Román Rivera and Karina Valarezo González

Estimating Customers’ Profitability: Influence of RFM Attributes, Web Metrics and Product Data 293
 Sunčica Rogić, Ljiljana Kaščelan, and Vladimir Đurišić

Marketing, Geomarketing and IOT

Toward Consumer Sustainable Consumption: Examining Factors Influencing Green Product Purchase Intention 307
 Marleen Prigita and Yeshika Alversia

A Smart City IoT Crowdsensing System Based on Data Streaming Architecture 319
 Aleksandra Labus, Miloš Radenković, Stefan Nešković, Snežana Popović, and Svetlana Mitrović

IoT Application Model in Secondary Education 329
 Danijela Stojanović, Zorica Bogdanović, and Marijana Despotović-Zrakić

Digital Marketing Strategy: A Step-By-Step Approach 339
 M. J. Guerra and F. O. Silva

Non-invasive Individual Sensing System for Collecting Biometric Indicators 351
 Francisco Monteiro, José Martins, Ramiro Gonçalves, and Frederico Branco

Machine Learning Applied to Marketing

Image Processing: Impact of Train and Test Sizes on Custom Image Recognition Algorithms 365
Luis Marques, Luca Lopes, Miguel Ferreira, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi

Neuromarketing Technologies

Demand Analysis for the Creation of a Technology Application to Improve Decision Making 383
Micaela Silva-Barragán and Carlos Ramos-Galarza

Neuromarketing: Current Applications in Favor of Consumerism 389
María Judith López and Carlos Ramos-Galarza

Expectations versus Perceptions: The Role of Consumers’ Senses in Composing of USP for PDO Wines 395
Jiří Zelený, Karolina Macháčková, Jan Hán, and Michal Lošťák

Enterprise Information System for Digitalization and Organization of Innovation—A Case Study 407
Zornitsa Yordanova

What Changed in One Year of a Pandemic and What the Portuguese are not Willing to Admit: Consumer Neuroscience and Predictive Analytic Contributes to Communication Strategy 419
Valentina Chkoniya, Dorota Reykowska, Rafal Ohme, and Ana Côrte-Real

Pre-ordering Music Vynils: The Impact of Price and Number of Band Supporters 431
Sónia Felizardo, Pedro Bem-Haja, Beatriz Casais, Sónia Brito-Costa, and Hugo de Almeida

Technologies Applied to Tourism Marketing

360° Technology Applied to Touristic Marketing 443
Omar Córdor-Herrera and Carlos Ramos-Galarza

Digital Marketing and User-Generated Content: A Case Study of Vidago Palace Hotel 451
Irina Clara, Teresa Paiva, and Elisabete Paulo Morais

Websites Usability Evaluation of the Terras De Trás-Os-Montes Hotels 463
Elisabete Paulo Morais, Carlos Rompante Cunha, and Arlindo Santos

Quality Factors for Agro-touristic Websites—An Exploratory Study	471
Karina Cicha and Paulina Rutecka	
Tracking Techniques in the Study of Tourists’ Spatiotemporal Behavior	485
Irina Clara, Ricardo Correia, and Aida Carvalho	
Organisation, Classification and Analysis of Online Reviews Directed to Retail in the Municipality of Porto	497
Pedro Braga, Pedro Quelhas Brito, and Mafalda Teles Roxo	
Strategic Digitization of Tourism	515
José J. Galán, Ramón A. Carrasco, and Antonio Latorre	
Tourism Professionals’ Awareness of the New Technologies: A Premise of the Destinations’ Modernization and Development	527
Diana Foris, Gabriela-Ramona Bâra, and Simona Mălăescu	
Social Media Usage Patterns and Motivations in Tourism Companies: An Outlook	539
Vasco Agostinho, Cristina Estevão, Paulo Duarte, and Gisela Gonçalves	
Digital Eco Tree Platform: A Proposal of an Effective Mobile Application in World Heritage Cities	551
Irena Travar, Zamira Acosta-Rubio, Eduardo Parra- López, and Vidina Tais Díaz-Padilla	
Virtual and Augmented Reality in Marketing	
An Integrated Framework of the Relation Between Augmented Reality and Brand Love	565
Carolina Afonso and Vanessa Hipólito	
A View of Augmented Reality in the Beauty Industry from an Exploratory Perspective: Generations X and Z	575
Sergio Barta, Raquel Gurrea, and Carlos Flavián	
Place Attachment Through Virtual Reality: A Comparative Study in Douro Region (Northern Portugal) with Video and ‘Real’ Visit	585
Nieves Losada, Filipa Jorge, Mário Sérgio Teixeira, Nuno Sousa, Miguel Melo, and Maximino Bessa	
The Use of Virtual Reality as an Advertising Tool for Brand Experience in Peru	595
Javier Espinoza-Nieves and Francisco Arbaiza	
Immersive Digital Marketing for Smart Cities Focusing Tourism	605
Carlos R. Cunha, Luisa Lopes, and Vítor Mendonça	

Digital Marketing and Branding

Are Colors Emotional Triggers in Digital Branding? 617
Joana Sampaio Correia and Dora Simões

The Brand-Cause Fit in the Advertising Campaign for Sprite’s #YouAreNotAlone 631
Claudia Cachay-Marín, Francisco Arbaiza, and Eliana Gallardo-Echenique

Consumer Bullying in Online Brand Communities—Quantifying a Dark Social Media Phenomenon 641
Oliver Warke, Jan Breitsohl, and Joemon Jose

Consumer Perception of Fast-Moving Consumer Goods About the Continente Food Lab Brand 653
João Pedro Carvalho and Ricardo Mena

Gender Role Stereotypes as an Ethical Resource in the Peruvian Advertising Discourse 665
Ninowska-Camila Carbajal-Obando, Melina Mezarina, and Eliana Gallardo-Echenique

Consumer Involvement in the Co-creation of a Disruptive Offer: The Case Study of *Continente Labs* 675
Clésia Varandas and Mafalda Nogueira

Brands at the Nexus of Organizations and Markets: How Organizational Structures and Technology Shape Marketing Practice 693
Miguel Magalhães and Fernando Pinto Santos

The Impact of Brand Community on the Consumer Behavior: The Harley-Davidson Brand Community in Portugal Case 703
Joaquim Pratas, Maria João Vasconcelos, and José Luís Reis

Luxury Fashion Brands’ Website Strategies: A Study with Portuguese Designers 717
Sandrina Teixeira, Belém Barbosa, Cátia Ferreira, and José Luís Reis

Correction to: Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution C1
Filomena Castro Lopes, Paula Morais, and Abílio Cardoso

Author Index 733

About the Editors

José Luís Reis done Ph.D. in Technologies and Information Systems from the University of Minho and is professor with the title of specialist in Management and Administration by IPAM—Porto. He is a professor at ISMAI and IPAM and an integrated researcher in LIACC—Laboratory of Artificial Intelligence and Informatics of the University of Porto. It carries out activities in the area of training and information systems and technologies in various organizations, coordinating various national and international projects in the area of information management, applied marketing, and strategic regional planning. He is the author of scientific papers and articles in the fields of information systems modeling, multimedia, gamification, and data mining. He is the author and co-author of several books, namely *Personalization in Marketing—Technologies and Information Systems*, *Marketing in Agri-Food—Fundamentals and Case Studies*, *Gamification Model for SMEs*, and *Marketing and Smart Technologies*.

Eduardo Parra López done Diploma in Business Studies, Degree in Economics and Business Administration, and Doctor from the University of La Laguna, Master International Business Program (EOI, 1997), and complemented with various courses of teaching and research specialization. He is the university professor of Business Organization in the Digital Economy and Tourism (University of La Laguna) and a former ex-president of the Spanish Association of Experts Tourism Scientists (AECIT, Period 2016–2020). It is worth highlighting his experience in the development of business plans, marketing and digital marketing for companies and public bodies in the tourism and commercial field, at the regional, national, and international. In the last 5 years (2015–2020), he has developed an important activity training, research, and outreach in Central America, which has allowed him a wide knowledge of the region's tourism ecosystem, as well as the operation of MSMEs tourism in Central America. As a result of his research activity in the 23 years of his university career, it is worth highlight the following: 60 publications, scientific documents in national journals and international impact indexes indexed in the main scientific impact indices and editions, authorship and co-authorship of books and technical documents; 61 participations in national and international congresses, as

well as 20 participations in scientific committees and congress organization; 26 organization of R + D + i activities and events and management of R + d + i (20). His experience in projects and institutions has focused mainly on the subject tourist and commercial. He is an external advisor for the Center American Tourism Agency (CATA). He has been visiting professor at the universities: University of Strathclyde (Scotland); IMI International Center for Tourism (Switzerland, Lucerne); Federal University of Parana (UFP, Curitiba); Del Valle University (Guatemala); National University of Nicaragua (UNA, Managua); Universidad Católica de la Paz (Bolivia); and University of Bristol (UK). Likewise, he is a visiting professor in Spain at the Universities of Vigo, Málaga, Oviedo, Almería, Lleida, and Valencia (Spain), in their master's programs in Tourism.

Luiz Moutinho held the Foundation Chair of Marketing at the Adam Smith Business School, University of Glasgow, Scotland, between 1996 and 2015. In 2017, Luiz Moutinho received a degree of Professor Honoris Causa from the University of Tourism and Management Skopje, Republic of Macedonia. Between 2015 and 2017, he held professorship of BioMarketing and Futures Research at DCU Business School, Dublin City University, Ireland. Actually, he is a visiting professor of Marketing at Suffolk Business School, University of Suffolk, Ipswich, England, and at the Marketing School, Portugal, he is adjunct professor at GSB, Faculty of Business and Economics, University of South Pacific, Suva, Fiji. He completed his Ph.D. at the University of Sheffield in 1982 and held posts at the Universidade Europeia, Lisbon, Portugal, the Hebei United University, Tangshan, China, Hubei University, China, Hubei University of Economics, China, University of the Highlands and Islands, Scotland, University of Coimbra, Portugal, University of Hull, England, Liverpool Hope University, England, University of Vilnius, Lithuania, University of Innsbruck, Austria, Otago University, New Zealand, and others. Professor Moutinho has over 172 articles published in refereed academic journals (Scopus), 35 books, and 17,125 academic citations, the h-index of 59 and the i10-index of 223 (Google Scholar, May 2021). Professor Moutinho is also a member of the Editorial Board of 52 international academic journals. He has also received several awards for excellence in academic research in the USA, UK, China, and Portugal.

José Paulo Marques dos Santos is an associate professor of Marketing and Qualitative Data Analysis at the University Institute of Maia (ISMAI), a visiting assistant professor of Neuromarketing and Qualitative Data Analysis at Universidade Europeia, and a researcher in the Unit of Experimental Biology of the Faculty of Medicine, University of Porto, and I3S—Instituto de Investigação e Inovação em Saúde, Portugal. Although originally a chemical engineer, José's professional life has been mostly devoted to management in the industry. After returning to academia, he entered a Ph.D. program on Neuroscience in Marketing, which concluded at the Technical University of Lisbon. His main interests are related to both these disciplines, emphasizing behavioral neuroscience, social neuroscience, and brands. Other interests include Semiotics, mainly Peircean Semiotics, and the processes of assigning meanings to symbols and the emergence of meanings within social

groups and the role of imitation in the social arena. In the methodology area, the interests are twofold: neuroimaging techniques as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), transcranial magnetic stimulation (TMS), transcranial direct/alternate current stimulation (tDCS/tACS), and functional near-infrared spectroscopy (fNIRS), together with multivariate analytical tools of the functional data like artificial neural networks (ANNs), and the construction of theory, with prominence to grounded theory, which is also a matter of concern and study.

Artificial Intelligence Applied in Marketing

Forecasting Hotel-booking Cancellations Using Personal Name Records: An Artificial Intelligence Approach



Eleazar C.-Sánchez , Agustín J. Sánchez-Medina ,
and Laura Romero-Domínguez 

Abstract Booking cancellation has a significant impact on hotel management and on the hospitality industry in general. According to some authors, an important portion of hotel revenue loss stems from not considering cancellations in booking systems. Besides, very little is still known about the reasons why customers cancel their bookings, or how this can be effectively avoided from happening. Keeping this in mind, personal name record (PNR) data, which were collected by a hotel partner, were used to design a model to forecast hotel-booking cancellations, which achieved 74% of accuracy. The benefits of the proposed method go beyond such good rate of detection: It only considers 13 different, independent variables, which is a reduced number compared to previous works in the field. Moreover, the included variables coincide with those requested from customers during the hotel reservation process. This is an advantage for hospitality establishments, given that these variables are often the only ones available. Our method allows knowing the cancellation rate with good accuracy, but it can also identify those customers who are likely to cancel their bookings. This approach could be an asset for organizations, as it assists them in improving their action protocols regarding incoming tourists.

Keywords Forecasting model · Hotel-booking cancellation · Support vector machine · Genetic algorithm · Personal name records · Machine learning

E. C.-Sánchez · A. J. Sánchez-Medina

Instituto Universitario de Ciencias y Tecnologías Cibernéticas (IUCTC), Universidad de Las Palmas de Gran Canaria, Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, Spain

e-mail: eleazar.caballero101@alu.ulpgc.es

A. J. Sánchez-Medina

e-mail: agustin.sanchez@ulpgc.es

L. Romero-Domínguez (✉)

Escuela Universitaria de Turismo de Santa Cruz de Tenerife, Calle José Zárate y Penichet 5, 38001 Santa Cruz de Tenerife, Spain

e-mail: laura.romero@eutur.es

1 Introduction

Tourism is one of the most important economic activities globally. Considering that the number of international arrivals increased from 400 million in 1990 to 1300 million in 2017, the growth and development of this industry are noteworthy. Regarding its economic contributions, tourism generates revenue through multiple ways such as consumption, taxes, business creation, and employment, among others. However, the tourist activity is characterized by the fact that it occurs in a highly uncertain context, i.e., there are multiple, uncontrollable factors which might affect tourist revenue, e.g., current political situation, climatological events [1]. In the particular case of the hospitality sector, since the product offered is intangible and cannot be stored, at the end of the day, each unsold room means lost income [2]. This reality requires a profound understanding of the short-term demand, along with the planning of the upcoming demand, in order to optimize hotel occupancy [2, 3]. Therefore, organization and planning are primordial, and forecasting highly contributes to such tasks [1]. According to Kourentzes et al. [4], demand forecasting is pivotal to organizations because it assists decision-making in multiple business areas. Nevertheless, predicting hotel demand is difficult [5, 6] due to its changing, unpredictable nature [7], which makes forecasting a challenging task [5]. Since hotel demand fluctuates, with bookings been possibly canceled, it is paramount for hotels to have a precise method to determine their actual occupancy, in order to assist decisions regarding revenue management [8]. It is evident that management decisions are affected by customer demand, which is determined not only by bookings, but also by cancellations [3, 9]. Concerning this matter, according to Sierag et al. [10], approximately 20% of accommodation revenue is lost because cancellations are not included in booking management systems. Despite its importance, the academic research on cancellation forecasting in the tourist sector is still scarce [9, 11, 12], and there is still much to learn about why guests cancel their bookings, and how to avoid this from happening [13].

Due to the above reasons, for the sake of the hospitality industry, the present work proposes and assesses an artificial intelligence (AI) model designed to forecast hotel-booking cancellations using personal name records (PNR). The use of PNR data constitutes a relatively novel approach in tourism [14], as well as the application of big data techniques to PRN data [15]. However, the existing literature on this field disagrees on the validity of PNR (cf. Romero Morales and Wang [16] and Antonio et al. [11, 17]). For this reason, this work also aims to elucidate the potential of PNR to predict hotel cancellations at the individual level.

2 Literature Review

2.1 *The Context of the Hospitality Industry*

Demand forecast and revenue management are closely related [18]. Due to the uncertain nature of the tourist environment, hotels must face many risks when managing room occupancy. In fact, since hospitality is such a demand-sensitive industry, unanticipated decreases in the demand have serious consequences for hotels [19]. Moreover, the uncertainty of the demand obviously affects occupancy management and planning, but it also influences other organizational matters, e.g., budgeting [20]. Now, demand might not be correctly predicted if cancellations are not considered. Thus, the net-demand, which responds to a “bookings minus cancellations” formula [16], should be taken into account [3]. To confront this situation, hotels implement a wide variety of actions to manage these risks (e.g., overbooking strategies, cancellation penalties). However, these measures might be counter-productive, because they might lead to revenue loss, rejection from customers [9], and negative effects on hotels’ corporate social reputation [11]. Consequently, having an accurate tool to predict cancellations becomes crucial [6], as it allows hotels to optimize occupancy and, at the same time, improve their strategies.

Now, there are some highlights that justify even more the relevance of reliable, precise cancellation prediction. With the development of IT, customer behavior has significantly changed, and forecasting their future demands and cancellations has become a complex issue. They are more informed about organizations and their services than they were decades ago (e.g., they compare offers by reading other customers’ experiences and opinions). Booking and canceling hotel services have also become much easier. Web sites allow customers to make several reservations in different establishments for a same period. Then, when they find the most suitable option, they select it and cancel the others [21]. As a result, the demand coming from Web sites increases, along with cancellations [11, 21, 22].

Besides, last-minute bookings have also contributed to raise the importance of net-demand forecasting. Last-minute bookings reduce the length of the reservation window, which affects the accuracy of predictions [23]. Additionally, they increase cancellations because customers who have already made a reservation, when encountering a last-minute, more economic offer, might cancel the former in favor of the latter. Thus, when last-minute offers are offered, the odds of cancellation increase as the service time approaches.

2.2 *Demand and Cancellation Forecasting: Methods and Techniques*

Pattern identification and forecasting are always based on previous experience with the service. However, the challenging aspect of such tasks is finding the most suitable way to carry them out, considering the availability of the data, as well as the established time and objectives [24]. In this section, we discuss the existing methods for demand and cancellation prediction, which have been applied in the hospitality sector. We distinguish between qualitative and quantitative techniques [25, 26].

Qualitative Techniques. These rely on a group of experts who, by means of available data and their self-expertise, determine tendencies and probabilities. Qualitative techniques should be used when the unprecedented nature of upcoming events does not allow appropriate forecasting, because historical data lack information (in both terms of quantity and quality) about said events. These could be, for instance, long-term predictions involving substantial changes [24] such as the popularization of ecotourism or short-term predictions involving unprecedented changes which are expected to cause a significant impact at the business level [27], e.g., the COVID-19 pandemic. Two outstanding qualitative techniques are the Delphi method and scenario writing [28].

In the particular case of the hospitality industry, Kaynak and Cavlek [29] carried out Delphi to predict the potential tourist market in Croatia. Alternatively, Lee et al. [24] forecast the number of visitors to the International Expo Tourism which took place in Korea in 2012. Lastly, the World Health Organization recently turned to an expert panel to forecast the recovery of international tourism after the COVID-19 pandemic [30].

Quantitative Techniques. To perform these techniques, sufficient and suitable historical data are needed [27]. If these data are quantifiable and past patterns can be inferred to the future, quantitative techniques are ideal [24]. In this subsection, we discuss the most important quantitative techniques for forecasting. These, following Peng et al. [31] and Song and Li [25], are grouped into non-causal time series-, econometric-, and AI-based models. It should be highlighted that most cases discussed in the following paragraphs focus on tourism destination forecasting, in contrast with hotel reservation forecasting, which is scarcer in the literature [32].

Non-causal time series models aim to uncover future patterns from historical data. Historically, the integrated autoregressive moving average (ARIMA) model has been the most widely used model to predict demand. However, in past years, seasonal ARIMA models (e.g., SARIMA) have popularized due to the close relationship between tourism and seasonality [25]. Concerning this matter, in order to analyze their role in the prediction of the tourist demand from four different European markets in Catalonia (Spain), Claveria and Datzira [33] included consumer expectations in time series models. These authors concluded that, among the different techniques applied, the ARIMA and the Markov switching regime (MKTAR) models showed the best performance, and those models including consumer expectations did not offer better outcomes in any of the time scenarios considered. Regarding the forecasting of

hotel demand, Pfeifer and Bodily [34] followed a space–time ARIMA (STARIMA) approach in order to predict arrivals in eight hotels, all of them belonging to the same chain and located in an American city. The authors concluded that STARIMA, which assumes dependence among points and gives a higher weight to the closer ones, performed better than a single ARIMA time series model.

Alternatively, among the econometric models, we can distinguish between static models, e.g., regression models, gravity models or the static almost-ideal demand system (AIDS), and dynamic models, e.g., vector autoregressive (VAR), time-varying parameters (TVP), or error correction model (ECM) [32]. An application of these models can be found in Falk and Vieru’s [35] work. These authors examined the factors influencing hotel-booking cancellations and considered variables such as length of stay, hotel category, and reservation time, among others. At a hotel level, they estimated the cancellation odds throughout a probit model with cluster-adjusted standard errors and discovered that, compared to offline and travel agency bookings, online reservations showed higher cancellation rates. Moreover, they also concluded that lead time during the booking process, as well as country of residence, had the highest effect, especially in the online option.

Lastly, AI-based techniques have been successfully applied to predict tourist and hotel demand [36]. Regarding cancellations, Romero Morales and Wang [16] used data mining to predict cancellation rates for service-booking revenue management. They considered 14 different variables, e.g., price, room category, booking channel. The authors stated that tree-based and kernel methods (particularly, support vector machine, SVM) were the most robust method to forecast hotel cancellations. Alternatively, Huang et al. [37] applied back propagation neural networking (BPNN) and general regression neural networking (GRNN) to forecast booking cancellations and concluded that both techniques adequately fulfilled said objective. Lastly, Antonio et al. [11] used several two-class classification algorithms (namely boosted decision trees, decision forests, decision jungles, locally deep support vector machine, and neural networks) to predict cancellation rates for four hotels located in the Algarve region (Portugal). The authors found that machine learning algorithms (particularly, decision forests) were adequate techniques to forecast hotel cancellations.

3 Methodology

The present work used actual, historical booking records to forecast future hotel cancellations. The data were obtained from a 4-star hotel partner located in Gran Canaria (Spain). To fulfill the research objectives, the extended CRISP-DM approach [38] was used.

3.1 Data Preparation

This research has been developed using personal name records (PNR). These sort of data are formed from the data which customers provide when placing a reservation, e.g., sale channel, number of guests, contracted services. In this study, approximately two years of booking records (from January 2016 to April 2018), which gathered more than 10,000 hotel bookings, were used. Provided PNR contained information from more than 30 different countries. Among all of them, Germany and United Kingdom gathered nearly half the total reservations. Ninety-seven percent of guests stayed in the hotel up to 14 days, being 7 days the most frequent length. Regarding cancelation behavior, approximately 30% of bookings were canceled before the scheduled day of arrival.

In order to design a useful hotel-cancelation forecast model, the most popular booking (e.g., booking, Trivago, Tripadvisor) and large hotel chain (e.g., Radisson, RIU, Meliá) Web sites were examined, with the aim of identifying the most common variables requested from customers when placing a reservation. The variables shown in Table 1 were found in all Web sites reviewed. In order to prevent a lack of natural variation, those items with missing values were discarded. Additionally, as we intended to optimize time and computational resources, accessing individual records were not deemed necessary.

When performing this kind of research, both number and type of variables are of paramount importance. Therefore, new variables can be generated from the original

Table 1 Explanatory variables included in the model

Name	Description	Type
Status	Reservation status (confirmed/canceled)	C
Adults	Number of adults	M
Company	Company through which the guest booked their room (e.g., TUI, Expedia)	C
Nationality	Guest's nationality	C
Advance payment	Advance payment required/not required	C
Length of stay	Number of nights spent at the hotel	M
Notice period	Difference between reservation and arrival date	M
Creation day	Day when the reservation was registered	M
Creation month	Month when the reservation was registered	M
Check-in day	Effective check-in day	M
Check-in month	Effective check-in month	M
Mean price	Average price of the room	M
Sales channel	Sales channels used for the reservation (e.g., B2B, hotel Web site)	C
Weekend days	Number of Saturdays and Sundays during the stay	M

C categorical; *M* metric

dataset, e.g., weekend days, which were calculated based on both check-in and check-out dates. Besides, rows with missing values were discarded, and only closed bookings were considered, because there was no way of confirming whether the booking was effectively consumed. In this research, the target variable was the reservation status. Thus, it was treated as a binary variable with two possible outcomes: canceled and not canceled. As regards the other variables, given that they had different ranges, predictors were normalized to avoid model sensitivity and to keep consistency.

3.2 *Models and Validation*

We used the R statistical software [39] to apply two supervised learning algorithms, namely support vector machine (SVM) [40] and genetic algorithm (GA) [41]. SVM aims to find a hyperplane (or boundary) that optimally separates the points of one class from those from another class [16] by applying the principle of structure risk minimization. For two linearly separable classes, SVM attempts to find the resulting optimal hyperplane with the lowest generalization error by maximizing the margin, which is, according to Bishop, “the smallest distance between the decision boundary and any of the samples” [42]. Nonetheless, linearly separable classes are not usual in real life. As a solution, a kernel function is used to transform a given data space into a n -dimensional space [43, 44] where the classes become linearly separable [45]. For its part, the GA is an evolutionary algorithm which looks for the best solution possible for a given solution. It starts by generating a random population, and then, it performs multiple reproduction, crossover, and mutation activities, in order to find the solution that better fits the function [46].

As regards model validation, in contrast with extant research applying the k -fold technique [11, 16], in this work, the model’s performance was validated throughout repeated random subsampling, which is advisable when the outcome variable is binary and raw data are not balanced [47]. This methodology consists in dividing the dataset into a training set (which is used to build the model) and a testing set (which is used for to validate the model). This process is repeated multiple times, and in the end, the mean value is considered for the final assessment. As there is no limit in the number of runs (which is contrary to k -fold), this technique avoids reliability problems, which may occur when running training and testing only once. At the same time, by using repeated random subsampling, some observations may not be tested, and others may be tested more than once.

In repeated random subsampling, data are split into training- and testing sets to fit the model’s parameters and to do an unbiased assessment, respectively. The data used in this study were unbalanced, which might generate a poor training set if one of the classes has a reduced number of examples (because this leads to a model that ignores one of the classes). To avoid this from happening, although the overall cancellation rate was approximately 30%, we balanced both datasets, so each of them included 50% of each class. This procedure was repeated 100 times. Every time, different, randomly selected training and testing sets were used. Thus, the models were trained

and tested once every run. Along the whole process, both the results obtained by each method, as well as actual values (those which were intended to predict) were saved, so the two lists could be compared to evaluate performance.

4 Results for Cancellation Forecasting

In the following section, we start by presenting performance measures for the technique used (SVM-GA optimized), and then, we discuss the results of the model.

The most common method to analyze the results from a classification method is the confusion matrix. This is a labeled crosstab layout that reports the differences between the actual class and the predicted class, as forecast by the testing set [48]. To analyze the relationship between these two classes, several performance measures are used: recall, specificity, accuracy, precision, and F-score. Recall, which is also known as sensitivity, represents the number of actual true positives identified by the test. Specificity is the same metric as recall, but it refers to the number of actual true negatives [49]. Accuracy measures the proportion of cases that the model predicts correctly (i.e., true positive and true negatives). Precision represents the rate of true positives over the total of positives predicted. Finally, following Guneş et al. [50], the F-score measures “the distinction between two classes with real values” by combining the precision and recall metrics. Additionally, a simple method to analyze prediction results is the ROC curve (Fig. 1). This is a graph that visually shows the relationship between true positives and true negatives, as forecast by the model, which allows to assess the performance of the model [48].

The results of the aforementioned performance measures are shown in Table 2.

As it can be observed, SVM performs well. The achieved accuracy is 73%, and recall and specificity are balanced. At the same time, the precision measure (true positives/all positives) also shows a good result. Finally, the F-score, which is the harmonic means of precision and recall, shows the correct balance between these two metrics.

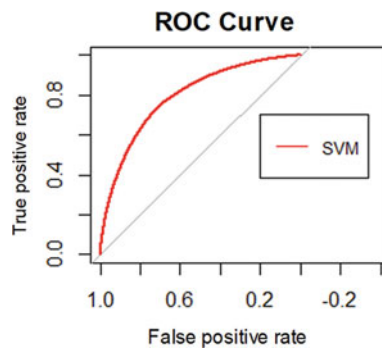


Fig. 1 ROC curve

Table 2 Performance measures

Method	Recall	Specificity	Accuracy	Precision	F-score	AUC
SVM (GA-optimized)	0.742	0.708	0.725	0.718	0.730	0.725

5 Conclusions

The present work contributes to broaden the limited extant literature on the hospitality sector by designing and validating an AI-based method to predict hotel-booking cancellations. Concerning its theoretical contributions, we have clarified that PNR data are indeed useful to predict individual cancellations, as suggested by Antonio et al. [11] and in contrast to Romero Morales and Wang [16]. Moreover, we have shown that individual cancellations can be predicted accurately with a reduced number of variables (13 in total). This is a significantly lower number than in previous research [11, 17].

From a practical perspective, another contribution of this work is the proposed methodology, as it allows accurate predictions with easily accessible variables. In this regard, our methodology uses variables which can be found at popular booking and hotel chain Web sites, which facilitates generating the database and avoids three important issues: extra queries, data on guests' identities [11], and external information sources [17]. Thus, we propose a simpler, faster forecasting methodology that allows a closer alignment of the predictions to the latest market trends (i.e., placing multiple bookings and canceling all of them except one, last-minute offers).

From a managerial perspective, it can be concluded that guests' historical data are a key resource for hospitality firms. Prediction tools might be used to treat such data and to assist hotel managers in decision-making. In other words, they may constitute an organizational advantage. Forecasting techniques allows hoteliers to have information on advance, which allows them to establish suitable overbooking and cancellation policies, to design adequate pricing strategies, etc. At the same time, the proposed methodology could turn into a competitive advantage, as it is capable to predict cancellations with high accuracy. However, it can also predict whether the customer is likely to cancel their booking. This would allow managers to anticipate events and avoid cancellations by, for instance, sending reminders to customers or directly contacting with them. They could even offer special gifts (e.g., chargeable services-free) depending on whether the guest is profitable enough. Similarly, individual cancellation policies could be individually established according to the cancellation odds (e.g., customers who are likely to cancel must pay a penalty to cancel their booking).

As future lines of research, we propose to test this methodology with different PNR data (e.g., hotel location, hotel category), so additional variables might be included in the model to improve its results. This could be done by adding extra fields to the booking forms (e.g., stay purpose—leisure or business) or by turning to other external resources (e.g., climate prediction, economic indexes from guests' countries of origin).

As research limitations are concerned, forecasting stems from historical data. Therefore, significant events and changes in the market could not be initially considered in the model.

References

1. Chow, W.S., Shyu, J.-C., Wang, K.-C.: Developing a forecast system for hotel occupancy rate using integrated ARIMA models. *J. Hosp. Tour. Manag.* **1**, 55–80 (1998). https://doi.org/10.1300/J268v01n03_05
2. Frechtling, D.C.: *Forecasting Tourism Demand: Methods and Strategies*. Butterworth-Heinemann, Oxford, Woburn (MA) (2001)
3. Rajopadhye, M., Ghalia, M.B., Wang, P.P., Baker, T., Eister, C.V.: Forecasting uncertain hotel room demand. *Inf. Sci.* **132**, 1–11 (2001). [https://doi.org/10.1016/S0020-0255\(00\)00082-7](https://doi.org/10.1016/S0020-0255(00)00082-7)
4. Kourentzes, N., Rostami-Tabar, B., Barrow, D.K.: Demand forecasting by temporal aggregation: using optimal or multiple aggregation levels? *J. Bus. Res.* **78**, 1–9 (2017). <https://doi.org/10.1016/j.jbusres.2017.04.016>
5. Li, S., Chen, T., Wang, L., Ming, C.: Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. *Tour. Manag.* **68**, 116–126 (2018). <https://doi.org/10.1016/j.tourman.2018.03.006>
6. Pan, B., Yang, Y.: Forecasting destination weekly hotel occupancy with big data. *J. Travel. Res.* **56**, 957–970 (2017). <https://doi.org/10.1177/0047287516669050>
7. van Ostaijen, T., Santos, B.F., Mitici, M.: Dynamic airline booking forecasting. In: *Proceedings of the 21st Air Transport Research Society World Conference*. Antwerp (2017)
8. Haensel, A., Koole, G.: Booking horizon forecasting with dynamic updating: a case study of hotel reservation data. *Int. J. Forecast.* **27**, 942–960 (2011). <https://doi.org/10.1016/j.ijforecast.2010.10.004>
9. Zakhary, A., Atiya, A.F., El-Shishiny, H., Gayar, N.E.: Forecasting hotel arrivals and occupancy using Monte Carlo simulation. *J. Revenue. Pricing. Manag.* **10**, 344–366 (2011). <https://doi.org/10.1057/rpm.2009.42>
10. Sierag, D.D., Koole, G.M., van der Mei, R.D., van der Rest, J.I., Zwart, B.: Revenue management under customer choice behaviour with cancellations and overbooking. *Eur. J. Oper. Res.* **246**, 170–185 (2015). <https://doi.org/10.1016/j.ejor.2015.04.014>
11. Antonio, N., de Almeida, A., Nunes, L.: Predicting hotel booking cancellations to decrease uncertainty and increase revenue. *Tour. Manag. Stud.* **13**, 25–39 (2017). <https://doi.org/10.18089/tms.2017.13203>
12. Antonio, N., de Almeida, A., Nunes, L.: Predictive models for hotel booking cancellation: a semi-automated analysis of the literature. *Tour. Manag. Stud.* **15**, 7–21 (2019). <https://doi.org/10.18089/tms.2019.15011>
13. Hajibaba, H., Boztuğ, Y., Dolnicar, S.: Preventing tourists from canceling in times of crises. *Ann. Tour. Res.* **60**, 48–62 (2016). <https://doi.org/10.1016/j.annals.2016.06.003>
14. Gorin, T., Brunger, W.G., White, M.M.: No-show forecasting: a blended cost-based, PNR-adjusted approach. *J. Revenue. Pricing. Manag.* **5**, 188–206 (2006). <https://doi.org/10.1057/palgrave.rpm.5160039>
15. Pereira, L.N.: An introduction to helpful forecasting methods for hotel revenue management. *Int. J. Hosp. Manag.* **58**, 13–23 (2016). <https://doi.org/10.1016/j.ijhm.2016.07.003>
16. Romero Morales, D., Wang, J.: Forecasting cancellation rates for services booking revenue management using data mining. *Eur. J. Oper. Res.* **202**, 554–562 (2010). <https://doi.org/10.1016/j.ejor.2009.06.006>
17. Antonio, N., de Almeida, A., Nunes, L.: Big data in hotel revenue management: exploring cancellation drivers to gain insights into booking cancellation behavior. *Cornell. Hosp. Q.* **60**, 298–319 (2019). <https://doi.org/10.1177/1938965519851466>

18. Tse, T.S.M., Poon, Y.T.: Analyzing the use of an advance booking curve in forecasting hotel reservations. *J. Travel. Tour. Mark.* **32**, 852–869 (2015). <https://doi.org/10.1080/10548408.2015.1063826>
19. Yüksel, S.: An integrated forecasting approach for hotels. In: *International Symposium on the Analytic Hierarchy Process (ISAHP) 2005*, 10. Honolulu (2005). <https://doi.org/10.13033/isahp.y2005.014>
20. Tang, C.M.F., King, B., Pratt, S.: Predicting hotel occupancies with public data. *Tour. Econ.* **23**, 1096–1113 (2016). <https://doi.org/10.1177/1354816616666670>
21. Chen, C.C., Schwartz, Z., Vargas, P.: The search for the best deal: how hotel cancellation policies affect the search and booking decisions of deal-seeking customers. *Int. J. Hosp. Manag.* **30**, 129–135 (2011). <https://doi.org/10.1016/j.ijhm.2010.03.010>
22. Mayr, T., Zins, A.H.: Acceptance of online vs. traditional travel agencies. *Anatolia* **20**, 164–177 (2009). <https://doi.org/10.1080/13032917.2009.10518902>
23. MacCarthy, B.L., Blome, C., Olhager, J., Srari, J.S., Zhao, X.: Supply chain evolution—theory, concepts and science. *Int. J. Oper. Prod. Manag.* **36**, 1696–1718 (2016). <https://doi.org/10.1108/IJOPM-02-2016-0080>
24. Lee, C.K., Song, H.J., Mjelde, J.W.: The forecasting of International Expo tourism using quantitative and qualitative techniques. *Tour. Manag.* **29**, 1084–1098 (2008). <https://doi.org/10.1016/j.tourman.2008.02.007>
25. Song, H., Li, G.: Tourism demand modelling and forecasting—a review of recent research. *Tour. Manag.* **29**, 203–220 (2008). <https://doi.org/10.1016/j.tourman.2007.07.016>
26. Song, H., Qiu, R.T.R., Park, J.: A review of research on tourism demand forecasting: launching the annals of tourism research curated collection on tourism demand forecasting. *Ann. Tour. Res.* **75**, 338–362 (2019). <https://doi.org/10.1016/j.annals.2018.12.001>
27. Uysal, M., Crompton, J.L.: An overview of approaches used to forecast tourism demand. *J. Travel. Res.* **23**, 7–15 (1985). <https://doi.org/10.1177/004728758502300402>
28. Lin, V.S., Song, H.: A review of Delphi forecasting research in tourism. *Curr. Issues. Tour.* **18**, 1099–1131 (2015). <https://doi.org/10.1080/13683500.2014.967187>
29. Kaynak, E., Cavlek, N.: Measurement of tourism market potential of Croatia by use of Delphi qualitative research technique. *J. East–West Bus.* **12**, 105–123 (2007). https://doi.org/10.1300/J097v12n04_05
30. World Health Organization (2021) 2020: Worst year in tourism history with 1 billion fewer international arrivals. <https://www.unwto.org/news/2020-worst-year-in-tourism-history-with-1-billion-fewer-international-arrivals>
31. Peng, B., Song, H., Crouch, G.I.: A meta-analysis of international tourism demand forecasting and implications for practice. *Tour. Manag.* **45**, 181–193 (2014). <https://doi.org/10.1016/j.tourman.2014.04.005>
32. Lee, M.: Modeling and forecasting hotel room demand based on advance booking information. *Tour. Manag.* **66**, 62–71 (2018). <https://doi.org/10.1016/j.tourman.2017.11.004>
33. Claveria, O., Datzira, J.: Forecasting tourism demand using consumer expectations. *Tour. Rev.* **65**, 18–36 (2010). <https://doi.org/10.1108/16605371011040889>
34. Pfeifer, P.E., Bodily, S.E.: A test of space-time arma modelling and forecasting of hotel data. *J. Forecast* **9**, 255–272 (1990). <https://doi.org/10.1002/for.3980090305>
35. Falk, M., Vieru, M.: Modelling the cancellation behaviour of hotel guests. *Int. J. Contemp. Hosp. Manag.* **30**, 3100–3116 (2018). <https://doi.org/10.1108/IJCHM-08-2017-0509>
36. Wu, D.C., Song, H., Shen, S.: New developments in tourism and hotel demand modeling and forecasting. *Int. J. Contemp. Hosp. Manag.* **29**, 507–529 (2017). <https://doi.org/10.1108/IJCHM-05-2015-0249>
37. Huang, H., Chang, A.Y., Ho, C.: Using artificial neural networks to establish a customer-cancellation prediction model. *Prz. Elektrotech* **2013**, 178–180 (2013)
38. Huber, S., Wiemer, H., Schneider, D., Ihlenfeldt, S.: DMME: data mining methodology for engineering applications—a holistic extension to the CRISP-DM model. *Procedia CIRP* **79**, 403–408 (2019). <https://doi.org/10.1016/j.procir.2019.02.106>

39. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (2020). <http://www.r-project.org>
40. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chih-Chung, C., Chih-Chen, L.: Package ‘e1071’ (2019). <https://CRAN.R-project.org/package=e1071>
41. Scrucca, L.: GA: a package for genetic algorithms in R. *J. Stat. Softw.* **53**, 1–37 (2013). <https://doi.org/10.18637/jss.v053.i04>
42. Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer, New York (2006)
43. Campbell, W.M., Sturim, D.E., Reynolds, D.A.: Support vector machines using GMM super-vectors for speaker verification. *IEEE Signal Process Lett.* **13**, 308–311 (2006). <https://doi.org/10.1109/LSP.2006.870086>
44. James, G., Witten, D., Hastie, T., Tibshirani, R.: *An Introduction to Statistical Learning: With Applications in R*. Springer, New York (2013)
45. Amari, S., Wu, S.: Improving support vector machine classifiers by modifying kernel functions. *Neural Netw.* **12**, 783–789 (1999). [https://doi.org/10.1016/S0893-6080\(99\)00032-5](https://doi.org/10.1016/S0893-6080(99)00032-5)
46. Murali, R.V., Puri, A.B., Prabhakaran, G.: GA-driven ANN model for worker assignment into virtual manufacturing cells. In: *Proceedings of the World Congress on Engineering 2010*. Newswood Limited, Hong Kong (2010)
47. Khakifirooz, M., Chien, C.F., Chen, Y.-J.: Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower industry 4.0. *Appl. Soft. Comput.* **68**, 990–999 (2018). <https://doi.org/10.1016/j.asoc.2017.11.034>
48. Bradley, A.P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern. Recognit.* **30**, 1145–1159 (1997). [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
49. Altman, D.G., Bland, J.M.: Diagnostic tests 1: sensitivity and specificity. *BMJ* **308**, 6943 (1994). <https://doi.org/10.1136/bmj.308.6943.1552>
50. Güneş, S., Polat, K., Yosunkaya, Ş: Multi-class f-score feature selection approach to classification of obstructive sleep apnea syndrome. *Expert Syst. Appl.* **37**, 998–1004 (2010). <https://doi.org/10.1016/j.eswa.2009.05.075>

Managerial Decisions in Marketing: The Individual Perception of Explainable Artificial Intelligence



Gioia V. Volkmar

Abstract This paper examines how marketing managers' perceived explanation quality differs between Artificial Intelligence (AI) and human recommendations. The results of this experimental study extend previous studies on managerial decisions by demonstrating that participants' perception of the explanation quality differed between AI and humans. The explanation behind human advice was perceived as less satisfying, needing more information about the recommendation. In contrast, receiving an AI-generated advice positively changed their perception of the explanation quality. They perceived the explanation as more sufficient without the need for further clarification. These findings contribute to understand the individual perception of the explanation quality of AI-generated recommendations. Including explainability of AI in the decision process is important for building user trust. Thus, this paper is intended to stimulate further research on the use of AI for managerial decisions in marketing, particularly examining the importance of explanations in the decision process to improve AI acceptance.

Keywords Marketing management · Decision-making · Explainable artificial intelligence

1 Introduction

Artificial intelligence (AI) has revolutionized countless fields over the past years and is constantly evolving [1]. AI is not only being applied in several different consumer applications but also within various industries, such as financial services, health care, commerce, and marketing [2]. Its greatest potential will be in enhancing human abilities. This will enable a collaboration between humans and machines that is able to solve various problems. Only through this collaboration trust can be established [3].

G. V. Volkmar (✉)
University of St. Gallen, 9000 St. Gallen, Switzerland
e-mail: gioia.volkmar@unisg.ch

There are not only benefits associated with AI, but also a wide criticism is connected with the risks of implementing AI. Stephen Hawking stated that it is still unclear whether AI can be seen as the best or worst thing for humanity [4]. This demonstrates the conflicting attitudes surrounding AI. On the one hand, researchers are questioning the ethicality of enhancing human capabilities. On the other hand, there is unquestionable potential of using AI in various business contexts, such as for managerial decisions [5]. However, it is critical for marketing managers to not only blindly rely on AI decisions, especially when the result of a decision is undesirable. In particular, one dimension plays an important role to build and retain trust in AI—decision explicability. According to EU Commission (2019), a certain explicability should be embedded in the decision-making process in order to ensure transparency [6]. Thus, future research is required to evaluate the ethical dimension and minimizing the negative perception of AI to assure favorable developments for the next years.

Based on prior research, an experiment has been conducted with marketing managers to find out whether the perceived explicability differs between AI and human decision input. This paper hence combines the theoretical background with the findings of the experimental study, derives implications, and shows possibilities for future research.

2 Theoretical Background

A growing number of companies are discussing possibilities of implementing AI into their organizations [7]. While AI is said to have the potential to improve efficiency and productivity within organizations [8], it is still being faced with increasing levels of skepticism whether AI technologies will transform the functioning of the workforce or should be seen as threat [9].

In the past years, AI has shown huge potential within decision-making [10]. Especially in marketing, the potential of implementing AI is becoming quite fruitful due to the possibilities of increasingly affordable and accessible data [11]. However, combining these two fields shows that marketing managers lack to tap the potential of AI technologies [12].

A core issue when interacting with AI is managing and maintaining trust in its decision-making process [13]. Managers tend to work on highly classified subjects, and thus, a lack of trust will most likely hinder them to use AI without their supervision [7]. The lack of transparency of AI systems makes it further difficult to understand the reasoning behind AI-based decisions [3]. In addition, the output of AI is fully based on its input, making the data sets used very crucial for a desired outcome. If it picks up on a biased view, AI could perform well for the wrong reasons [14]. Therefore, including transparency in the decision-making process is crucial to successfully use AI [3]. Thus, a certain level of explainability is important to incorporate in the AI reasoning in an effort to maximize transparency [15]. Explanations can be seen as making a word, an idea, or a concept understandable to others [16].

When looking on prior research, two prominent and yet divergent bodies of literature can be distinguished: One research stream states that people prefer human recommendations since they are easier to understand. If an algorithmic recommendation was presented with a certain explanation, individuals would be less averse of using it. In this case, the explainability and understandability are key rather than a recommendation system's ability to outperform humans [17]. Another research stream discusses that lay people prefer advice from AI compared to human recommendations [18]. Such individuals tend to perceive advice from an expert system to be more rational and objective than human advice [19]. Thus, blindly trusting recommendations without closer scrutiny or questioning the advice can be very dangerous. Doing so hence may result in failing to notice errors in an expert system [20].

Most research on AI still focuses predominantly on the transformation of unstructured and structured data into Smart Data (e.g., [21]) or implementing Smart Technologies into consumer environments (e.g., [22, 23]). However, little research investigates how managers perceive a certain explainability of AI-based vs. human decision input.

3 Experimental Study: Research Procedure and Results

The experimental study tested the explicability of managerial decisions in marketing, taken either by AI or by humans, by understanding the decision attributes involved. This study will investigate how far explicability plays a role in the decision process.

3.1 Research Procedure

Data Collection and Sample

An online experiment was conducted with 250 managers ($M_{\text{age}} = 40$, $SD = 10.04$). As this paper focuses on managerial decision-making, managers formed the population of interest. In order to assure a marketing management sample, the participants were eliminated without any employee responsibility. Their attention was tested by adding attention check items, such as asking about their company business [24]. The effective sample size was thus reduced and resulted in $n = 207$.

Method

A scenario-based approach was chosen for the experiment, using a common decision-making situation in marketing, more specifically in product development. In this way, conclusions about real-life behavior can be drawn [25]. The second part of the study started with follow-up questions regarding the experiment. In order to measure the perceived explanation quality, the item from Dijkstra (1999) regarding potential questioning of the advice was adapted [19]. Further, the items based on Stalmeier

et al. (2005) to evaluate a decision regarding their perception of an informed decision and satisfaction of the explanation were modified [26].

3.2 Results

First, the relationship between the source of advice (human vs. AI) and the perceived explanation quality (no explanation was given in both conditions) was tested. The regression results show that the source of advice significantly affects the perceived explanation quality of the advice ($\beta = 0.75$, $t = 2.78$, $p < 0.01$). Participants, who received a human recommendation, perceived the explanation as less satisfying and good compared to an AI-generated advice.

Second, in order to understand the effect of the source of advice on the perceived explanation quality, the decision evaluation (perceived informed decision) was included as a mediator. The bootstrapping method based on 5.000 bootstrap samples [27] was used to test the relationship between the source of advice and perceived explanation quality through decision evaluation. The source of advice had a significant effect on the decision evaluation ($b = 0.47$, $SE = 0.17$, $p < 0.005$), which had a significant effect on the perceived explanation quality ($b = 1.10$, $SE = 0.08$, $p < 0.0001$). After including decision evaluation as a mediator, the direct effect between the source of advice on the explanation quality became nonsignificant ($b = 0.23$, $SE = 0.21$, $p > 0.10$). Thus, the decision evaluation fully mediated the impact of the source of advice on the perceived explanation quality. Further mediation analysis based on the bootstrapping method indicates that the effect of the source of advice on the perceived explanation quality is mediated by decision evaluation ($b = -0.52$, 95% CI [0.17, 0.87]). This means that in the human condition, participants felt less informed about the received recommendation and thus were not satisfied with the explanation. In comparison, when receiving an AI-generated advice, they felt more informed and hence perceived the explanation as having better quality.

4 Implications and Future Research

This study aims to initiate further research on the use of AI for managerial decisions in marketing, especially the importance of explainability in the decision process. First, the results show that even though no explanation was given, participant's perception of the explanation quality differed between both conditions (human vs. AI). The explanation of the human advice was perceived as less satisfying, and they felt the needed more information about the recommendation. In contrast, receiving an AI-generated advice significantly changed their perception of the explanation quality. Managers perceived the explanation as more sufficient without the need to ask further

questions. Second, the results demonstrate the mediating role of the decision evaluation (perceived informed choice). When receiving the AI-generated recommendation, the participants felt more informed and were more satisfied with the explanation which positively affected the perceived explanation quality. In contrast, this effect was significantly weaker with human advice. Participants wanted more information about the recommendation and thus rated the explanation quality as inferior.

The results of the experimental study have important implications for practitioners and researchers. *For practitioners*, since marketing managers are required to both take and deal with many strategic and highly classified decisions, a certain transparency is crucial to ensure a desired decision outcome. Thus, it is not sufficient to blindly trust AI recommendations, but to include a certain degree of explainability of AI-generated recommendations. *For researchers*, these results make important contributions to the literature on managerial decisions, AI, and psychology. In particular, it shows the differing perception of explainability in the decision process depending on the source of advice. These individual impressions should be actively used in order to increase the acceptance of AI.

This paper serves as starting point for further studies and thus has further research potential. *First*, since the recommendation was presented without context, participants did not receive an explanation. Future research could manipulate this by providing different levels of explainability. *Second*, additional measurements of explainability should be incorporated in the next studies. These scales could include how objective, logical, relevant, reliable the explanation was perceived between the two conditions. These measurements can help to further examine the reasons underlying the results presented in this study. *Third*, future studies could include different levels of urgency or impact of the decision outcome and analyze whether there are any changes in the results.

5 Conclusion

This paper aims to explore how marketing managers perceive the explanation quality and how it differs between AI and human-based recommendations. Therefore, the results of this experimental study seek to contribute to the existing literature on managerial decision-making in marketing, psychology, and AI technology. Until now, there has been no comparable study that successfully links these three topics and examines the difference between the source of advice (human vs. AI) and the perceived explanation quality. It is hence important to understand the individual perception of the explanation quality of AI-generated recommendations.

The findings of the experimental study provide evidence that the perception of the recommendation differs when generated by AI or humans and should play an important role in the decision-making process. These results can lead to various consequences; hence, it is crucial that managers be made aware of these when it comes to the acceptance of AI.

References

1. Torra, V. et al.: Artificial intelligence, In: Said, A., Torra, V. (eds.) *Data science in practice*, Chapter 2, 9–26, Springer (2019)
2. Huang, M.-H., Roland, T., Rust, R.: Artificial intelligence in service. *J. Serv. Res.* **21**(2), 155–172 (2018)
3. Jarrahi, H.: Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. *Bus. Horiz.* **61**, 577–586 (2018)
4. Saffell, N.: “The best or worst thing to happen to humanity”—Stephen Hawking launches Centre for the Future of Intelligence. Retrieved from <https://www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of>. Last accessed 02 Aug 2021
5. Edwards, J.S., Duan, Y., Robins, P.C.: An analysis of expert systems for business decision making at different levels and in different roles. *Eur. J. Inf. Syst.* **9**(1), 36–46 (2000)
6. EU Commission: Ethics Guidelines for trustworthy AI. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Last accessed 12 Jul 2021
7. Kolbjørnsrud, V., Richard Amico, R., T. Robert J.: How Artificial Intelligence will redefine Management. *Harvard Bus. Rev.*, 1–6 (2016)
8. Paschen, J., Kietzmann, J., Kietzmann, T. C.: Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. *J. Bus. Industrial Market.*, 1410–1419 (2019)
9. Daugherty, P., Wilson, J.: Collaborative intelligence: humans and AI are joining forces. *Harv. Bus. Rev.* **96**(4), 114–123 (2018)
10. Duan, Y., Edwards, J.S., Dwivedi, Y.K.: Artificial intelligence for decision making in the era of Big Data—evolution, challenges and research agenda. *Int. J. Inf. Manage.* **48**, 63–71 (2019)
11. Conick, H.: The past, present and future of AI in marketing. *Marketing News*. <https://www.ama.org/marketing-news/the-past-present-and-future-of-ai-in-marketing/>. Last accessed 05 Jul 2021
12. Davenport, T., Guha, A., Grewal, D., Bressgott, T.: How artificial intelligence will change the future of marketing. *J. Acad. Mark. Sci.* **48**(1), 24–42 (2020)
13. Davenport, T.H., Kirby, J.: Just how smart are smart machines? *MIT Sloan Manag. Rev.* **57**(3), 21 (2016)
14. Ransbotham, S., Kiron, D., Gerbert, P., Reeves, M.: Reshaping business with artificial intelligence: closing the gap between ambition and action. *MIT Sloan Manag. Rev.*, 59(1) (2017)
15. Brewer, W.F., Chinn, C.A., Samarapungavan, A.: Explanation in scientists and children. *Mind. Mach.* **8**(1), 119–136 (1998)
16. Kim, J.: Explanation. In: R. Audi, R. (ed.): *The Cambridge dictionary of philosophy*, 298–299, Cambridge, England, Cambridge University Press (1995).
17. Yeomans, M., Shah, A., Mullainathan, S., Kleinberg, J.: Making sense of recommendations. *J. Behav. Decis. Mak.* **32**(4), 403–414 (2019)
18. Logg, J.M., Minson, J.A., Moore, D.A.: Algorithm appreciation: people prefer algorithmic to human judgment. *Organ. Behav. Hum. Decis. Process.* **151**, 90–103 (2019)
19. Dijkstra, J.: User agreement with incorrect expert system advice. *Behaviour Inf. Technol.* **18**(6), 399–411 (1999)
20. Dijkstra, J., Liebrand, W.B., Timminga, E.: Persuasiveness of expert systems. *Behaviour Inf. Technol.* **17**(3), 155–163 (1998)
21. Wedel, M., Kannan, P.: Marketing analytics for data-rich environments. *J. Mark.* **80**(6), 97–121 (2016)
22. Novak, T., Hoffman, D.: Relationship journeys in the internet of things: a new framework for understanding interactions between consumers and smart objects. *J. Acad. Mark. Sci.* **47**, 216–237 (2019)
23. Kietzmann, J., Paschen, J., Treen, E.: Artificial intelligence in advertising. How marketers can leverage artificial intelligence along the consumer journey. *J. Advertis. Res.* **58**(3), 263–267 (2018)

24. Oppenheimer, D.M., Meyvis, T., Davidenko, N.: Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* **45**, 867–872 (2009)
25. Woods, S., Walters, M., Koay, K., & Dautenhahn, K.: Comparing human robot interaction scenarios using live and video-based methods: towards a novel methodological approach. *Proceedings of 9th IEEE international workshop on advanced motion control*, 750–755 (2006)
26. Stalmeier, P.F., Roosmalen, M.S., Verhoef, L.C., Hoekstra-Weebers, J.E., Oosterwijk, J.C., Moog, U., van Daal, W.A.: The decision evaluation scales. *Patient Educ. Couns.* **57**(3), 286–293 (2005)
27. Hayes, A. F.: *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*. Guilford Publications (2013)

Find Me if You Can! Identification of Services on Websites by Human Beings and Artificial Intelligence



Christian Stadlmann , Gábor Berend , Alexandra Fratrič ,
Zsolt Szántó , David Aufreiter , Margarethe Ueberwimmer ,
and Stefan Mang

Abstract This paper presents the results of evaluating the number of service offerings marketed on company webpages in the field of advanced manufacturing performed by human annotators, artificial intelligence imitating human behavior, and advanced in-depth AI analysis. The research focuses on ten different countries and three specific sectors: manufacture of computer, electronic and optical product, manufacture of electrical equipment, and manufacture of machinery and equipment. Even though artificial intelligence was able to find more services on company webpages than human beings, the average number of services identified on webpages proved to be relatively low. Companies from selected industries focus mostly on product lifecycle services, such as spare parts, repair services, or maintenance. On the other hand, the least marketed services are in the field of finances, such as pay per use and installment payment. The results support the applicability and effectiveness of artificial intelligence tools in the field of procurement. Moreover, the findings are highly relevant for companies in the field of advanced manufacturing as they indicate a great potential for further improvement of service promotion on company webpages.

Keywords Artificial intelligence · Servitization · Manufacturing · Service promotion

1 Introduction

Companies and scholars have realized the impact of the Internet on the purchasing process. Almost 75% of buyers in the B2B segment prefer searching and selecting goods and services online instead of communicating and engaging with a salesperson

C. Stadlmann · A. Fratrič · D. Aufreiter (✉) · M. Ueberwimmer
University of Applied Science Upper Austria, Wehrgrabengasse 1–3, 4400 Steyr, Austria
e-mail: david.aufreiter@fh-steyr.at

G. Berend · Z. Szántó
University of Szeged, Árpád tér 2, 6720 Szeged, Hungary

S. Mang
University of Passau, Innstraße 41, 94032 Passau, Germany

[1, 2]. Moreover, as purchasing is an integral part of successful business operations and can provide significant value, researchers and company representatives look for ways to utilize new technology to optimize the process as well as to find suitable suppliers and service providers [3, 4].

As buyers commonly browse the Internet to find suitable offerings and companies move their presence online, it is imperative that suppliers promote their products and service offerings on their webpages and social media.

2 Promoting Services on Company Webpages

As global markets have been changing rapidly and became highly competitive, complex, and often unpredictable, companies are looking for new ways to ensure their survival and profitability [5–7]. Current literature points to a servitization trend among manufacturers, where combining product offerings with services is perceived as a strategic alternative to developing new innovative products [6, 8–13]. Yielding considerable advantages, such as more stable revenue streams, improved profitability, or increased customer satisfaction and loyalty, services help companies to differentiate and create a sustainable source of competitive advantage [14]. With the rapid rate of new technology adoption in organizations, the focus in the context of servitization has been shifting toward the potential and benefits of utilizing the Internet and digital technologies in business processes [15, 16]. Companies have moved their presence and marketing activities online and procurement personnel commonly browses company webpages to obtain information about service characteristics and prices and to search for the most suitable offer [17, 18].

Buyers increasingly rely on information they find on the Internet while identifying new business opportunities, conducting market research, or preliminary identification of suppliers [19, 20]. The most significant benefits of online research in procurement are the ability to compare products, services, and prices, the speed, and ease of use (finding information they need effortlessly), and convenience [21, 22]. Moreover, reviewing product or service information online increases the accuracy of goods purchased [23].

Manufacturing companies can take advantage of these procurement trends as their websites might reach customer segments, which might not have been possible to target in the past [24]. Proper design of webpages in line with company market positioning enables to promote key quality elements and establish the reputation—especially in B2B service segments [24, 25]. Literature suggests that having a company webpage is linked to higher revenues due to increased familiarity with products and services offered [26]. However, in order to exploit the potential presented by company webpages, certain requirements have to be fulfilled. Aspects determining the user intention to use webpages during the procurement process are the accuracy of information provided on the webpage, how easy information can be found, and buyers' perception of the usefulness of the webpage [22, 27]. Trustworthy presentation of information enhances perceived value [18]. The main aim of the webpage

design should be customer satisfaction, which can be increased by providing detailed product or service information in a simple yet attractive format. Service providers have to adequately signal on their webpage what services they offer and that they are capable of facilitating high-quality services [25]. Therefore, this paper focusses on how intensively manufacturing companies promote services on their company webpages [28]. To obtain this evaluation of service promotion on webpages, various methods of webpage analysis were employed to fully understand not only if services are offered on company webpages but also if they are offered in a manner that enables human beings to find them.

Hence, this article examines and answers the following research questions:

- How intensively are services promoted on webpages in the field of advanced manufacturing?
- What are the differences in service research on company webpages performed by humans, AI solutions imitating human behavior, and in-depth analysis using improved AI algorithms?

3 Methodology

3.1 *Empirical Context—Service Typology in the Field of Advanced Manufacturing*

The European Union defines advanced manufacturing as “the use of knowledge and innovative technology to produce complex products [...] and improve processes to lower waste, pollution, material consumption and energy use.” [29] Important elements are artificial intelligence, robotics, 3D and 4D printing, and high-performance computing for modeling. The industry is an important segment of the economy and labor market in the European Union, with 14.5 million jobs and a strategic focus in many regional and national strategies [29–31].

The research presented in this article focuses on the webpages of companies in the field of advanced manufacturing in three specific sectors: manufacture of computer, electronic and optical product (NACE 26), manufacture of electrical equipment (NACE 27), and manufacture of machinery and equipment (NACE 28). To identify and structure services, which would be analyzed on company webpages, the research consortium adjusted the service typology presented by Partanen [32]. The main service categories utilized in this research are pre-sales services, product support services, product lifecycle services, as well as R&D, operational, and financial services. A detailed overview of the categories assessed can be found in Table 1.

As the research activities were a part of an Interreg Central Europe project called “ProsperAMnet” financed by the European Union Development Fund, the research focused on and was conducted for the following countries: Austria, Germany, Czech Republic, Hungary, Italy, Slovakia, Slovenia, France, United Kingdom, and the

United States of America. The lists of companies from each country were obtained from national and international databases such as Europages, Aurelia, and others. Only companies operating in the specified sectors (NACE 26, 27, 28) were analyzed. To ensure accurate representation of companies within each industry, companies with more than 20 employees were selected regardless of whether they operate nationally or internationally.

3.2 Applied Methods for Analyzing Webpages

In order to be able to assess the degree of service promotion on company webpages in the field of advanced manufacturing as well as to identify differences in various methods of company webpage service search, the consortium developed a research process comprising of three steps. Firstly, the traditional market research was performed using human annotators to investigate selected company webpages (see part 3.3). The data set gathered in this step also formed the basis for the training of the developed artificial intelligence algorithms. This was accomplished by manually extracting specific information from company webpages in selected industries. To standardize the process, a scheme to perform the annotations as well as a structured template was designed to extract and classify the data. In the second research phase, the artificial intelligence algorithm was developed in order to imitate human behavior and large-scale automated webpage research was executed. In the third step, an advanced artificial intelligence algorithm was created and implemented to in-depth analyze service offerings on company webpages. The final artificial intelligence algorithm should be able to answer the general question: “Which companies promote which kind of services on their webpages?”. In the final stage of the research, the results from all three methods were compared.

3.3 Human Behavior: Manual Annotations

In total 1.487 company webpages were manually analyzed through annotations between March and September 2020 with the objectives to understand the webpage promotion of services and to generate training data for AI algorithms. The analysis was conducted in English, German, French, Italian, Slovak, Slovenian, Hungarian, and Czech languages by twenty-two native speakers.

In the first method, data from company webpages was manually extracted according to a comprehensive structure, categorized, and subsequently analyzed. In order to collect the data in a structured manner, a detailed manual was designed. The researchers from seven different countries were carefully selected and trained. They were graduate or postgraduate students from the field of business and/or technology, had domain knowledge, and were native speakers of the languages specified in the text above (with the only exception being the English language). The scheme

used for the analysis consisted of specific information, which was to be retrieved from company webpages—company name, URL, country, language, location of headquarters, number of employees, NACE code, identified services from the list of services (Table 1).

In order to standardize the process and minimize the discrepancies between different researchers, a trial process of annotation was applied. After all, researchers collected company information according to the predefined structure, the data was cross-checked, differences among the results were identified, and comprehensive feedback was presented to the researchers in order to improve the manual research process. This process was repeated in three loops to ensure high quality of the data collection.

For the quantitative measurement of the inter-annotator agreement, we had 15 cases when the same company website was annotated by a pair of independent annotators. The annotation process was open-ended, i.e., annotators were not provided with a fixed set of websites from a given URL to check but a base URL of a company was given to them and then they were allowed to check any of the in-domain webpages on the given URL. As such, a pair of annotators could find evidence for a particular service being offered by some company on two distinct pages of the same company website. In order to account for that, we aggregated the services found by the individual annotators at the company level, i.e., the result of each annotation with respect to a company website consisted of binary values, each one indicating whether the annotator deemed a service to be offered based on the contents of a company website.

We calculated the Jaccard coefficient over the 15 pairs of annotations, more specifically, we took the fraction of the aggregated number of such services that were identified in consensus by a pair of annotators and that of such services that were found to be relevant by at least one of the annotators [33]. Besides this empirical quotient, we also quantified the expected value of the Jaccard coefficient, more specifically, the quantity that we would obtain if the annotators were choosing the same number of services as they did during the annotation, except for the fact that they chose the selected number of services randomly. Assuming that a pair of annotators identified m and n services for a given company website out of S potential services being distinguished, the expected value of their Jaccard coefficient can be given as

$$\frac{nm}{S_n + S_m - nm}$$

The observed Jaccard coefficient ended up 0.377, whereas the expected value calculated as described above was 0.143. A typical way to contrast observed (o) and expected (e) values of random variables are to calculate the fraction $\frac{o-e}{1-e}$, resulting in 0.272 in our particular case. This value is not particularly high, which illustrates the complexity of the annotation process. This value, however, also indicates that the annotators agreed to a non-trivial extent.

Twenty-two researchers manually extracted data from 1487 company webpages; 613 annotations were performed by the Austrian researchers, 193 by the Hungarian,

198 by the German, 198 by the Italian, 97 by the Slovenian, 94 by the Slovak, and 94 by Czech investigators. Depending of the amount on information presented, each company webpage analysis took on average 20 min to complete.

Discrepancies among the investigators were overcome and limited due to the quality loop controls, which standardized the research process and the interpretation methods. During the investigation process, differences among the webpages regarding their professionalism were identified. Moreover, certain service categories (such as “customer seminars” or “technical user training”) were not sufficiently distinguished on the webpages by companies. These hurdles were diminished through the domain knowledge of the researchers and peer-reviews, discussions, and interpretations of these challenging webpages or service presentations. Furthermore, a critical mass of company webpages with higher content quality was guaranteed in this process, in order to generate the basis for the second method and to prevent and limit vagueness of information provided.

3.4 Web Analysis Using Artificial Intelligence

In the second step, the data obtained from the manual webpage analysis was used as training data to perform identification of offered services on webpages. Based on the annotations, we trained a logistic regression classifier, which assigns a weight for every combination of words from the training data and service categories investigated. A higher weight indicated that some word was more likely to serve as evidence for a particular kind of service being offered.

The input features of our logistic regression model were derived from the individual words included on the analyzed websites. In order to overcome the limited size of the training data and to be able to process websites in multiple languages, we replaced every content word found on the websites with their five most similar English synonyms/translations.

In order to do so, we performed an automatic language detection on the contents of the websites and relied on cross-lingual word embeddings (CLWEs) to find the most likely synonym/translation of a word [34]. Word embeddings assign a vector to every word form in a way that words with similar meaning get assigned to vectors with similar orientations. CLWEs also have the desirable property, that they are insensitive to the particular language a word is written in, e.g., it assigns a similar vector to the German word ‘*Hund*’ and the English word ‘*dog*’, as they have the same meaning. We used our logistic regression classifier for the individual webpages of a company website traversing the websites in a priority queue. The priority queue contained at most 200 websites that could be reached via at most clicking three hyperlinks from the start page. The priority queue contained the URLs ranked by an additional model which tries to predict the likelihood that the given URL contained some evidence being offered by the company based on the URL itself, e.g., an URL containing `/our-services/` is more likely to contain mentions of services a priori compared to another website that contains `/contacts/`. The limit for checking up at most

200 websites per company website at most 3 hops away from the starting page was introduced for performance reasons and also as human annotations generally behaved like that as well. As these choices resembled that of the annotators, we refer to this variant of our approach as our model imitating human behavior.

We also made a more permissive variant of our approach which also considered the analysis of those webpages that were up to 10 clicks away from the initial website to start the analysis from. We refer to this kind of analysis as the in-depth analysis hereinafter.

The logistic regression classifier can assign a probability value of every webpage promoting a certain service. For a company website we calculate all these probability values for all the services towards all the (at most 200) webpages our ranking module selects for analysis, then for each service category choose the highest probability value obtained and take it as the probability of the given company offering that given service. We conclude that a given service is offered by a company whenever the corresponding probability exceeds 0.5, i.e., it is more likely to be offered than not.

4 Findings

The results in Table 1 show considerable differences between the website research performed by human annotators and by the AI algorithm imitating human behavior. Even greater disparity can be observed when comparing human annotations with the results provided by the in-depth AI analysis. In all the service categories analyzed, the AI was more likely to find service offerings on company webpages compared to the annotations performed by humans. These discrepancies might be the result of ineffective and imprecise promotion of services on the webpages resulting in lower visibility of offerings for human researchers.

The service category, which is promoted by companies on the webpages at the highest rate is product lifecycle services. Within this service group, humans found the most service offerings in spare parts category (26.03%), whereas artificial intelligence detected the highest percentage of services in maintenance (38.87%). On the contrary, the category with the lowest rate of service offerings as identified both by humans and the AI is the financial services, more specifically pay per use services, where human annotations identified that only 0.2% of companies offer these services, compared to 0.94% identified by the in-depth AI analysis.

Table 2 shows that the average number of service offerings identified on a single webpage analyzed by humans is 2.12, compared to the results of 4.01 and 4.41 found by human-imitating AI and in-depth AI analysis respectively. Even though manufacturing is considered a sector with a high rate of servitization compared to other industries, all three figures show a relatively low number of services offered by companies, indicating a need for improvement and increase in service promotion in the industry.

Table 1 Results of service research on company webpages based on the service typology developed by Partanen et al. [32]

No	Service category	Human annotations		Prediction (AI imitating human behavior)		Prediction (AI imitating human behavior)		Prediction (in-depth AI analysis)		Prediction (in-depth AI analysis)	
		#	%	#	%	#	%	#	%	#	%
	Sample size	$n = 1487$		$n = 1487$		$n = 7740$		$n = 1487$		$n = 7740$	
	Number (%) of identified services	#	%	#	%	#	%	#	%	#	%
1	<i>Pre-sales services</i>										
11	Product demonstrations	41	2.76	141	9.48	694	8.97	168	11.30	821	10.61
12	Customer seminars	57	3.83	199	13.38	1125	14.53	244	16.41	1340	17.31
2	<i>Product support services</i>										
21	Warranty	157	10.56	343	23.07	1133	14.64	376	25.29	1295	16.73
22	Technical user training	288	19.37	415	27.91	1493	19.29	459	30.87	1670	21.58
23	Customer consulting and support by phone	251	16.88	460	30.93	1427	18.44	489	32.89	1579	20.40
24	Testing, test rigs, quality assurance	135	9.08	322	21.65	1320	17.05	351	23.60	1437	18.57
3	<i>Product lifecycle services</i>										
31	Installation services	294	19.77	512	34.43	2172	28.06	546	36.72	2372	30.65
32	Repair service	335	22.53	527	35.44	2255	29.13	562	37.79	2460	31.78
33	Spare parts	387	26.03	513	34.50	1998	25.81	554	37.26	2213	28.59
34	Maintenance	380	25.55	548	36.85	2088	26.98	578	38.87	2257	29.16

(continued)

Table 1 (continued)

No	Service category	Human annotations		Prediction (AI imitating human behavior)		Prediction (AI imitating human behavior)		Prediction (in-depth AI analysis)		Prediction (in-depth AI analysis)	
		#	%	#	%	#	%	#	%	#	%
	Sample size	<i>n</i> = 1487		<i>n</i> = 1487		<i>n</i> = 7740		<i>n</i> = 1487		<i>n</i> = 7740	
	Number (%) of identified services										
35	Retrofit, modernization, upgrades	200	13.45	353	23.74	1204	15.56	379	25.49	1356	17.52
4	<i>R&D services</i>										
41	Research service	86	5.78	319	21.45	1433	18.51	354	23.81	1575	20.35
42	Prototype design and development	227	15.27	470	31.61	2365	30.56	510	34.30	2499	32.29
43	Feasibility studies	63	4.24	237	15.94	1309	16.91	280	18.83	1471	19.01
5	<i>Operational services</i>										
51	Project management	118	7.94	290	19.50	1664	21.50	327	21.99	1829	23.63
52	Service for operating the product for the customer	7	0.47	23	1.55	85	1.10	29	1.95	104	1.34

(continued)

Table 1 (continued)

No	Service category	Human annotations		Prediction (AI imitating human behavior)		Prediction (AI imitating human behavior)		Prediction (in-depth AI analysis)		Prediction (in-depth AI analysis)	
		#	%	#	%	#	%	#	%	#	%
	Sample size	<i>n</i> = 1487		<i>n</i> = 1487		<i>n</i> = 7740		<i>n</i> = 1487		<i>n</i> = 7740	
	Number (%) of identified services			#	%	#	%	#	%	#	%
53	Service for operating customer's processes	28	1.88	138	9.28	684	8.84	169	11.37	809	10.45
6	<i>Financial services</i>										
61	Pay per use	3	0.20	9	0.61	34	0.44	14	0.94	52	0.67
62	Installment payment	16	1.08	25	1.68	77	0.99	31	2.08	93	1.20
63	Leasing	23	1.55	43	2.89	108	1.40	50	3.36	137	1.77
64	Rental system	56	3.77	74	4.98	299	3.86	88	5.92	349	4.51

Table 2 The average number of services identified on company webpages based on different research methods applied

Service category	Human annotations	Prediction (AI imitating human behavior)	Prediction (AI imitating human behavior)	Prediction (in-depth AI analysis)	Prediction (in-depth AI analysis)
Sample size	$n = 1487$	$n = 1487$	$n = 7740$	$n = 1487$	$n = 7740$
Average number of services identified on company webpages	2.12	4.01	3.23	4.41	3.58

5 Outlook and Limitations

The article contributed to literature in a relatively new and rapidly developing field. However, further research and verification are advised, such as a qualitative investigation analyzing the reasons of underrepresentation of services. Moreover, further quantitative research could provide insight and empirical understanding of companies' point of view and verify the results. This research also highlights the importance and necessity of action in the field of management and marketing of services, as the lack of service offerings on company webpages or their poor visibility pose a significant challenge for procurement performed by human employees. The results presented support the applicability and effectiveness of artificial intelligence tools in the field of procurement due to the low quality of information provided on webpages of suppliers. Further research should also establish if the lack of service offerings on company webpages stems from insufficient marketing or if companies are indeed not offering many services and are, therefore, not presented on webpages.

Certain biases might have affected the results of the article. Insufficient or noisy annotations by humans might have led to skewed results, which is particularly important for services with large gaps between the different stages of the research process. Annotators might have had different labeling preferences and styles, which could have resulted in discrepancies of information collected. Moreover, a sampling bias might have occurred due to the selection of particular types of companies.

Acknowledgements The work described in this document has been conducted as part of a project "ProsperAMnet", which is supported by the Interreg CENTRAL EUROPE Programme funded under the European Regional Development Fund. The developed tool described in the paper is provided at the domain www.prosperamnet.eu free for use.

References

1. Koponen, J.P., Rytsy, S.: Social presence and e-commerce B2B chat functions. *EJM* (2020)
2. Arli, D., Bauer, C., Palmatier, R.W.: Relational selling. Past, present and future. *Industrial Marketing Management* (2018)
3. Wynstra, J.Y.F.: Past, present and future trends of purchasing and supply management: an extensive literature review. *A Journey Through Manufacturing and Supply Chain* (199–228) (2016)
4. Pulles, N.J., Veldman, J., Schiele, H.: Identifying innovative suppliers in business networks: an empirical study. *Ind. Mark. Manage.* **43**(3), 409–418 (2014)
5. Anderson, J.C., Narus, J.A.: Capturing the value of supplementary services. *Harv. Bus. Rev.* **73**, 75–83 (1995)
6. Vandermerwe, S., Rada, J.: Servitization of business. Adding value by adding services. *Euro. Manag. J.* (1988)
7. Dimache, A., Roche, T.: A decision methodology to support servitisation of manufacturing. *Int. J. Op Prod. Manag.* (2013)
8. Carlborg, P., Kindström, D., Kowalkowski, C.: The evolution of service innovation research. A critical review and synthesis. *Service Industries J.* (2013)
9. Baines, T.S., Lightfoot, H.W., Evans, S., Neely, A., Greenough, R., Peppard, J., Roy, R., Shehab, E., Braganza, A., Tiwari, A., Alcock, J.R., Angus, J.P., Bastl, M., Cousens, A., Irving, P., Johnson, M., Kingston, J., Lockett, H., Martinez, V., Michele, P., Tranfield, D., Walton, I.M., Wilson, H.: State-of-the-art in product-service systems. *Proc. Inst. Mech. Eng., Part B: J. Eng. Manuf.* (2007)
10. Bustinza, O.F., Bigdeli, A.Z., Baines, T., Elliot, C.: Servitization and competitive advantage. the importance of organizational structure and value chain position. *Res. Technol. Manage.* (2015)
11. Parida, V., Rönningberg Sjödin, D., Wincent, J., Kohtamäki, M.: Mastering the transition to product-service provision: insights into business models, learning activities, and capabilities. *Res. Technol. Manage.* **57**(3), 44–52 (2014)
12. Kamal, M.M., Sivarajah, U., Bigdeli, A.Z., Missi, F., Koliousis, Y.: Servitization implementation in the manufacturing organisations. Classification of strategies, definitions, benefits and challenges. *Int. J. Inf. Manage.* (2020)
13. Rabetino, R., Kohtamäki, M., Brax, S.A., Sihvonen, J.: The tribes in the field of servitization. Discovering latent streams across 30 years of research. *Industrial Market. Manage.* (2021)
14. Guo, A., Li, Y., Zuo, Z., Chen, G.: Influence of organizational elements on manufacturing firms' service-enhancement. An empirical study based on Chinese ICT industry. *Technol. Soc.* (2015)
15. Coreynen, W., Matthyssens, P., van Bockhaven, W.: Boosting servitization through digitization. Pathways and dynamic resource configurations for manufacturers. *Industrial Market. Manage.* (2017)
16. Gago, D., Rubalcaba, L.: Innovation and ICT in service firms. Towards a multidimensional approach for impact assessment. *J. Evol. Econ.* (2007)
17. Bienhaus, F., Haddud, A.: Procurement 4.0. Factors influencing the digitisation of procurement and supply chains. *BPMJ* (2018).
18. Chang, Y.-Y., Lin, S.-C., Yen, D.C., Hung, J.-W.: The trust model of enterprise purchasing for B2B e-marketplaces. *Comput. Standards Interfaces* (2020)
19. Konuk, F.A.: The role of store image, perceived quality, trust and perceived value in predicting consumers' purchase intentions towards organic private label food. *J. Retail. Consumer Services* (2018)
20. Davila, T., Gupta, M.R., Palmer, R.J.: Moving procurement systems to the internet. The adoption and use of e-procurement technology models. *SSRN J.* (2002)
21. Bridges, E., Goldsmith, R.E., Hofacker, C.F.: Attracting and retaining online buyers. In: Clarke, I., Flaherty, T.B. (eds.) *Advances in Electronic Marketing*, pp. 1–27. IGI Global (2005)

22. Merrilees, B., Fenech, T.: From catalog to Web. B2B multi-channel marketing strategy. *Industrial Market. Manage.* (2007)
23. Vaidyanathan, G., Devaraj, S.: The role of quality in e-procurement performance. An empirical analysis. *J. Oper. Manage.* (2008)
24. Williams, Z., Lueg, J.E., Hancock, T., Goffnett, S.P.: Positioning through B2B carrier signals. Understanding how service quality is communicated via websites. *Industrial Market. Manage.* (2019)
25. Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R.: Signaling theory. A review and assessment. *J. Manage.* (2011)
26. Liu, C., Arnett, K.P., Capella, L.M., Beatty, R.C.: Web sites of the Fortune 500 companies. Facing customers through home pages. *Inf. Manage.* (1997)
27. Zhang, Y., Fang, Y., Wei, K.-K., Ramsey, E., McCole, P., Chen, H.: Repurchase intention in B2C e-commerce—a relationship quality perspective. *Inf. Manage.* (2011)
28. Liu-Thompkins, Y.: A decade of online advertising research. What we learned and what we need to know. *J. Advertis.* (2019)
29. European Commission: Why the EU supports advanced manufacturing. https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/key-enabling-technologies/advanced-manufacturing_en (2021). Accessed 21 Jul 2021
30. European Commission: Horizon 2020. <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/nanotechnologies-advanced-materials-advanced-manufacturing-and-processing-and> (2020). Accessed 21 Jul 2021
31. European Commission: Regional Innovation Monitor Plus. <https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/news/challenges-ris3-implementation-and-advanced-manufacturing-region-bratislava> (2021). Accessed 21 Jul 2021
32. Partanen, J., Kohtamäki, M., Parida, V., Wincent, J.: Developing and validating a multi-dimensional scale for operationalizing industrial service offering. *JBIM* (2017)
33. Deza, M.M., Deza, E.: *Encyclopedia of Distances*. Springer, London (2016)
34. Joulin, A., Bojanowski, P., Mikolov, T., Jégou, H., Grave, E.: Loss in Translation. Learning Bilingual Word Mapping with a Retrieval Criterion. In: Riloff, E., Chiang, D., Hockenmaier, J., Tsujii, J.'i. (eds.) *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, 01-Oct-18–01-Nov-18, pp. 2979–2984. Association for Computational Linguistics, Stroudsburg, PA, USA (2018)

Under Which Conditions Are Humans Motivated to Delegate Tasks to AI? A Taxonomy on the Human Emotional State Driving the Motivation for AI Delegation



Anna Bouwer

Abstract The intensity of human–artificial intelligence (AI) interactions has been growing at a rapid pace. Research has acknowledged a simultaneous significant resistance on the part of users towards AI services on the one hand and a profound acceptance of AI solutions on the other. As a remedy for this ambiguity concerning AI delegation, the author takes the next step to explain the decisive relevant factors. This research introduces the concept of the human emotional state driving the motivation for AI-based task delegation. Precisely, the affective state, as a function of the two independent neurophysiological systems of valence and arousal, determines the motivation for AI delegation in the individual decision situation. The interplay between these two determinants of a human’s affective state results in a four-quadrant taxonomy on AI delegation. For instance, a combination of low arousal and negative valence results in an affective state, which motivates the human to decide in favour of AI delegation; an opposite emotional state of high arousal and positive valence yields a low incentive to apply AI services. The implications of the present research provide novel reasons for the presence and extension of AI services in the fulfilment of human tasks.

Keywords Artificial intelligence · Emotional state · Human decision-making · Technology adoption · Technology aversion

1 Introduction

The use of artificial intelligence (AI) to complete the daily tasks and duties of humans has been continually increasing in recent years [31], [60], [91]. Next to mastering simple requests such as recommending a movie on Netflix or an accommodation on Airbnb [133], “AI appreciation” is reflected in AI delegation for tasks that were

A. Bouwer (✉)

University of St. Gallen, Dufourstrasse 50, 9000 St. Gallen, Switzerland

e-mail: anna.bouwer@unisg.ch

Institute of Behavioral Science and Technology, Torstrasse 25, 9000 St. Gallen, Switzerland

formerly performed exclusively by humans. For instance, AI services have been transforming health care by providing automated diagnosis and treatment for patients [71]. Simultaneously, AI is capable of creating unique artistic output such as the portrait sold at the famous British auction house Christie's [27] and self-driving cars are becoming increasingly prevalent [92]. Despite this progress in AI applicability, general anxiety concerning non-human intelligence, fear about superintelligence, and concerns regarding AI power abuse [116] have nourished what is referred to as "AI aversion" [53]. These situational decisions in favour of or against AI services have raised the question on the driving factors behind "AI appreciation" versus "AI aversion". Specifically, under which conditions and to what extent are humans motivated to delegate a task to AI?

This paper aims to develop a taxonomy that portrays that the motivation to delegate a task to AI is based on the emotional state of the user or human. In the first step, the author provides an introduction concerning the current state of research on AI appreciation and aversion, focusing on how AI augments and replaces human tasks and decisions specifically. The author discusses AI delegation in terms of (a) which tasks people are willing to give up and (b) why they are willing to give up these tasks. Next, the author creates a bridge to the impact of emotions on decision-making, suggesting that people only like to delegate certain tasks and that the specific discrete emotions associated with this task help explain delegation decisions. The author summarizes these findings using a taxonomy of AI delegation based on the emotional state of the human.

2 The Appreciation and Denial of AI Services

Resistance to rely on AI services has been investigated across a broad set of decision contexts and settings, from human resistance to medical AI [71] to the aversion to rely on forecasts made by a superior algorithm [36], [38], [40]. For instance, Longoni et al. [71] revealed that patients are less inclined to rely on AI across a broad range of medical decisions, from basic screening tasks to the selection of a treatment. These negative perceptions have been proven to persist even under conditions where the AI shows comparatively higher accuracy rates compared to human medical providers [71, 105]. Similar reservation towards AI has been found in the domain of self-driving automobiles [1], [56], [61], [98]. For instance, humans have been reported to fear that they will lose the fun and joy they experience while driving their car [33, 45, 61]. Feelings have been found to have considerable explanatory power for describing human attitudes towards self-driving cars [48, 106]. Despite this emotional deviation from AI, humans have been shown to be open to this type of technology innovation if they can maintain the control over the vehicle [2, 50, 117].

Research on human augmentation through AI has similarly suggested that even pairing AI and human input leads to negative spillover effects such that the same medical provider is judged as more negative compared to providers not relying on AI input [94, 122]. Castelo et al. [25] suggested that humans distrust, rely less on, and

reject AI input the more subjective as opposed to objective the task is. In summary, AI services are generally judged more negatively when the output is less predetermined [37], the environment is more non-standard [25], and when humans believe that the decision context is unique to their person [71].

Opposing these findings on AI aversion, researchers have assumed that superintelligence will be reached in the coming decades, with concerning impact on the human race [14, 20, 42, 85]. Moreover, experts have claimed the areas of AI applicability are developing and it is a mere question of time until humans will have to accept the omnipresent role of AI in their daily task delegations [46]. For instance, smart home appliances such as metres, grids, lightning systems, and kitchen products can be found with increasing presence in households [43, 59, 95, 103, 138]. In terms of device usage, the perceived benefits of smart home appliances, such as convenience, outweigh the perceived disadvantages [134] as these benefits are thought to empower the human in charge [58]. Overall, humans have been shown to apply smart home solutions as a means to solve daily household problems and these solutions have been found to be highly appreciated by the users [22].

Furthermore, humans rely increasingly on smart speaker appliances such as Google Home and Amazon Eco [57, 96]. For instance, humans request information, compose shopping lists, and complete other daily routines via voice-based commands [22]. Overall, the total usage of voice assistants adds up to over an hour every day in the US homes, during which 80% of the users search for real-time information, for instance, on weather or traffic [88]. The latest development has resulted from the COVID-19 pandemic, which resulted in a 5% increase in the quantity of daily smart speaker commands, especially due to orders from home [118]. For instance, the fast food delivery chain Domino's introduced the "Dom system", an AI voice assistant for placing phone orders [137], and expanded the options for a person to order food via smart home devices such as smart TVs [29]. Despite this increase in smart speaker appliances, researchers have found that humans are reluctant to use voice assistants to play games, buy products, or conduct online banking [22]. In this context, research has shown that behavioural intention regarding smart speakers is predominantly driven by enjoyment, which is a subjective, emotional argument on the part of the user [57]. Sohn and Kwon [127] confirmed this finding, labelling enjoyment as a critical determinant for human acceptance of AI-based products. Here, the authors highlighted that the significance of this emotional component driving human attitude is unique to AI applications.

Additionally, statistical-based deep learning recommender systems are growing in usage, such as in the retail industry [123]. For example, humans shopping via the e-commerce platform Amazon are making use of algorithms that propose future products based on variables such as past purchases, similar customer preferences, or demographics [64, 68]. Overall, the literature review of AI in service applications revealed a mixed attitude towards AI-based task completion. In certain scenarios, humans are more willing to make use of AI, but why? What are the driving forces towards embracement versus denial of AI-based solutions?

3 Emotions and Decision-Making

Psychologists have regarded emotion building as a complex, multidimensional process, which is not simply based on situational appraisal [41]. According to the latest research stream on emotion theory, emotions are the result of an ongoing conceptualization of a core affect [114] via the process of categorization [13]. For instance, internal sensory stimulation is categorized into a meaningful affective state [69] that, combined with other psychical conjunctions, comprises emotions [41]. Thus, affect can be utilized as a shared terminology to define feeling states such as emotions [89]. Precisely, emotions can be regarded as an interplay of two cognitive appraisal dimensions: arousal and valence [113]—two factors which form the foundation of the pleasure-arousal theory of emotions [109]. Here, arousal is measured by the extent the neural system is activated and valence by the degree to which an emotion is pleasant [104, 131]. These two independent neurophysiological systems [124] interrelate with each other to create a spatial model, which explains experienced affective states (i.e. emotional states).

Emotions fundamentally influence human decision-making [3, 8, 17, 52, 54, 75, 83, 101, 102, 119, 132, 139, 140]. For instance, the somatic marker hypothesis claimed that human decisions are largely based on the emotional quality of possible outcomes [17]. Moreover, emotions are not a static construct, but have been shown to change based on the stimulus the human is confronted with [87]. From a neuroscientific perspective, similar regions of the prefrontal cortex have been identified as modulating emotions as well as decision-making [16, 83]. A significant overlap in the neural region, called amygdala [102], has been found to be responsible for decision-making and emotion regulation [83, 121]. Furthermore, the amygdala has been identified as a contributor to arousal processing [55]. The stimulation of the anterior insula, which is involved in emotional processing [128], was demonstrated to forecast decision-making in the ultimatum game [132]. In contrast, a defect in neuroanatomical structures has an impact on emotion processing, which impairs decision-making [17].

4 Linking AI Delegation to Emotional States

The preceding sections have depicted that people only like to delegate certain tasks to AI and that emotions play a role in decision-making and delegation. Therefore, the author claims that understanding the relative emotionality triggered by a task can determine human motivation for AI-based task completion. Researchers confirm the significant role of emotions in explaining users' motivation to apply AI-based solutions. For instance, Lu et al. [73] labelled emotion as the foremost driver towards long-term motivation to utilize AI devices; contributors to technology acceptance research have confirmed this finding [44, 127]. Simultaneously, researchers have urged limiting of the reliance on the technology acceptance model [63] to

describe decision-making with respect to AI-based solutions [26, 44]. Thus, despite the existing research on the relationship between emotions and decision-making, researchers have constantly acknowledged that further, more specific knowledge on the interaction types and implications are needed [8, 97, 99, 126]. In the following section, the author proposes the following logic to explain human attitude towards and thus interaction degree with AI applications.

In the first step, a certain stimulus (i.e. the need to fulfil a task) reaches the human limbic system and activates the two independent neurophysiological systems: valence and arousal [78, 113]. Activation of the arousal dimension describes the degree to which the task is exciting or stimulating and activation of valence is the degree to which this task is perceived as enjoyable [104]. The combination of both results in an affective state [75, 109, 135], which has a distinct and systematic influence on decision-making [74, 75, 76]. Precisely, pleased subjects (i.e. positive valence) have been found to perceive decision tasks as positive experiences, whereas less pleased subjects (i.e. negative valence) perceive the task as a necessity rather than an enjoyment [76]. Further, the effect of processing amplitude is moderated by the subject's task involvement [77]. Precisely, greater involvement is found to activate the central path of cognitive processing (i.e. thoughtful consideration of arguments), whereas lesser involvement results in cognitive processing via the peripheral path (i.e. simple, affective heuristics) [100].

Comparable impact holds for the case of arousal concerning decision-making. Aroused subjects have been shown to focus their decision-making, possess a narrow attention span, and have an ability to prioritize key attributes in comparison with less aroused subjects [75]. Thus, high arousal has been found to result in polarization and judgement simplification [67, 70, 75] and thereby sets the boundaries for attentional capacity [76]. Overall, the combination of arousal and valence induces an inclination towards decision strategies in line with the human emotional state. In the following section, the author investigates the impact of arousal and valence on AI delegation.

5 A Taxonomy on Human Motivation to Delegate Tasks Towards AI

The following taxonomy explains the human attitude towards AI delegation in the decision situation at hand (i.e. the final step from affective state to a person's decision to delegate tasks to AI). Here, the interplay between valence (i.e. horizontal axis) and arousal (i.e. vertical axis) results in four quadrants (see Fig. 1). Each of the four quadrants describes whether humans are motivated to apply AI-based solutions. For instance, a human in the lower left corner is in favour of AI delegation, while humans in the upper right corner would refrain from AI delegation. Humans in the two remaining quadrants have a relative willingness to delegate a task to AI. Next, the author explains the underlying process for each of the four quadrants.

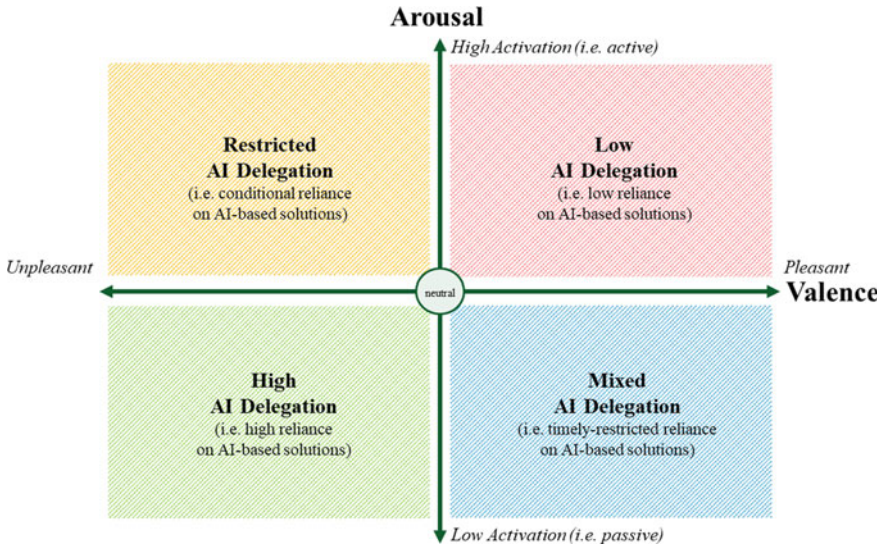


Fig. 1 Motivation for AI delegation based on the emotional state

In the lower left corner, which is named “**High AI Delegation**”, the person in charge is very willing to relinquish autonomy so that the task is conducted by an AI-based solution. Here, the task evokes a low amount of arousal and a feeling of unpleasantness for the person. The resulting emotional state combining low arousal and negative valence, such as feeling bored, favours the deliberate choice of AI use. Reviving the explanation of arousal concerning decision-making criteria, low arousal strengthens the focus of human decision-making on factors such as task objectivity, allowing for elaboration of the advantages of possible alternatives. Applying these observations in case of, for instance, autonomous domestic products implies that humans emphasize the relative advantage of AI-based solutions such as economic profitability or convenience. This example on emotion-based AI delegation taxonomy has been supported by the existing literature on autonomous domestic products [110]. Additionally, accounting for the negative valence in this quadrant, humans appreciate when “annoying” daily tasks are taken care of autonomously without their involvement [111]. Overall, it has been suggested that AI applications are favoured to fulfil tasks related to “tiresome” household chores [125], such as temperature regulation [112], lightning control [79], and smart kitchen applications [80].

Further, negative valence has been associated with low engagement [86]—a presage for missing identification with the task at hand. In turn, this missing identification has been demonstrated to drive the importance of functional features of autonomous technological applications [35]. This focus on functional features is strengthened by the low arousal, since only in case of low arousal are humans able to perform a thoughtful, in-depth, multi-perspective elaboration regarding the decision

at hand. Castelo et al. [25] confirmed the reliance on algorithms for presumably objective tasks. An example of such an objective task is driving a car, where the human is motivated to reach a target destination. For this person, driving a car is rather an annoying task that is neither activating nor pleasant [19]. Therefore, the elaborative thinking of the human, with a broad attention span in the decision situation, combined with the adverse associations regarding the task, results in AI appreciation. Thus, if a task is perceived as unpleasant and not activating, from the user's point of view, his or her willingness to complete a task by him or herself is low and the motivation to delegate to AI is high.

In the scenario of the upper right corner, the combination of high arousal and positive valence marks a “**Low AI Delegation**” affective state. Here, the task triggers the nervous system along valence and arousal, resulting in high activation and pleasantness. In this case, humans have been shown to be very unwilling to apply AI-based solutions (i.e. high motivation to complete the task by themselves), as they perceive the task as a positive experience, an enjoyment (i.e. high valence), which in turn drives the intrinsic motivation [108] for a self-accomplished activity. This intrinsic motivation is closely linked to identity-based consumption [107]. Precisely, humans perceive the utilization of technology-based solutions as fraud in cases where the task has an identity-based background [24, 66]. Accordingly, the sustainable, internal human motivation to pursue a task, paired with the desire to attribute task completion to oneself, drives the inclination not to forego this joyful task in favour of using AI. For example, a person who is passionate about automotive tasks enjoys every single second behind the wheel and is passionate about driving [12]. Therefore, this person is unlikely to make use of a self-driving-car, since he or she associates the task of driving with positive emotions, which in turn catalyse the person's internal motivation for self-attributed task completion. Leung et al. [66] confirmed that humans who strongly identify themselves with driving are more likely to own a car with a manual transmission.

The impact of arousal has been demonstrated to strengthen the inclination against task delegation to AI, since people enjoy high arousal [129] and joy is a direct attribute of intrinsic motivation [108], which in turn results in intrinsic satisfaction for self-accomplished tasks [81]. Precisely, a high level of arousal results in simplified and polarized decision-making. This implies that humans follow their initial, impulsive considerations regarding the task at hand (i.e. the joyful perception based on the high valence) and decide to fulfil the task by themselves. Thus, if a task is perceived as pleasant and activating from the user's point of view, his or her willingness to complete the task by him or herself will be high and the motivation to delegate to AI will be low.

In the lower right quadrant are decision situations where humans apply “**Mixed AI Delegation**” based on positive valence and low arousal. Precisely, the tasks that fall into this quadrant are pleasant but not stimulating to the person. The human is willing to make use of AI for mere activation reasons (i.e. raising arousal via AI-based solutions) until the point where arousal is high. Specifically, low arousal is associated with lesser task involvement [82], which results in low motivation to deal with the situation at hand [28] and therefore in high motivation to outsource initial

parts of the task to AI. Subsequently, the author claims that persons in the lower right quadrant will only make use of an AI application for intermediary tasks.

A comparable example is the use of e-commerce recommender systems, which endorse commercial products and services based on existing human data points [115, 136]. For example, humans have been making use of the Amazon algorithm, which proposes future products based on variables such as past purchases, similar customer preferences, or simple demographics [64, 68]. Hence, according to the taxonomy, humans will make use of the Amazon algorithm to preselect existing offers (i.e. a task where the arousal is very low), but execute the purchase by themselves (i.e. an emotionally exciting, alerting task). Aggarwal and Mazumdar [4] confirmed that the final purchase step is the most critical in terms of perceived loss in control, which results in an increased risk perception [18]. Riskiness, in turn, has been found to increase arousal (Schmidt et al., 2013). Subsequently, the initially favourable attributes of the AI solution such as enhanced decision quality and efficiency [5, 47] lose significance with increasing arousal, given increasing arousal shifts the narrow attention span away from the objective, well-elaborated advantages towards the polarized, subjective AI-associated risk factors [110]. Thus, once the task is not pre-selecting but deciding upon the actual purchase, the human steps in and wants to complete the arousing, positive final task by him or herself.

Further, the increasing amount of multimedia content over the past 10 years, and the resulting “information overload” [32], has led to extensive discussions surrounding affective recommender systems [84, 130], for instance, selecting a movie on Netflix; here, the company itself has confirmed that the ultimate goal is to retain joy for their members (i.e. create positive valence), implying that the Netflix Recommender System (NRS) [93] is applied to cover the, as the firm acknowledges, extensive search process for a relevant content [6]—a task with low arousal. Therefore, the person makes use of the NRS until the point where his or her arousal rises (i.e. the algorithm recommends an appealing movie). Thus, the person trades own task execution for AI delegation, but only for pre-selection reasons and not for the full task of selecting the movie.

In case the task is to purchase a product or service, which is unlikable and not stimulating (i.e. overall low involvement) for the human, he or she would be placed in the lower left quadrant (i.e. high AI delegation) according to his or her affective state. Accordingly, he or she would be very willing to make use of AI-based task completion (e.g. Alexa-based diaper purchases). This example shows that even when confronted with the same task, the human evaluation on AI applicability differs based on the personal judgement according to the affective state in the decision situation at hand.

The upper left quadrant describes the “**Restricted AI Delegation**” perspective on AI-based task completion. In this quadrant, the author positions all decision situations based on a highly activated but unpleasant affective state, decisions that raise emotions such as “stressed”, “nervous”, and “tense”. Thus, the upper left quadrant refers to decision situations with a high arousal and negative valence, whereby the latter implies that humans have negative associations with the task at hand [11] and perceive it as a burden, comparable to the lower left quadrant. Thus, the state of

valence stimulates the human to use AI. Despite this, humans in situations that are positioned in the upper left quadrant are especially hesitant to apply AI due to the high arousal in the situation at hand. A prominent example is the previously mentioned work of Longoni et al. [71] focused on medical AI, where the authors confronted the study participants with a decision regarding a medical provider for prevention of skin cancer. The majority of participants voted in favour of the dermatologist rather than the AI-based solution. This case exemplifies a situation of high arousal, which is triggered due to the confrontation with the AI raising human alertness as a reaction to the lowered trust (i.e. higher riskiness) in AI in comparison with a human professional [105]. High arousal is evident through increased heart rate and skin conductance levels [30], which are attributes that have been observed in uncomfortable medical situations [10, 49]. Recalling the fact that high arousal leads to polarized, simple decision-making, and is further increased with growing riskiness of the task [62], the result is a human inclination to refrain from AI as a yet unfamiliar [51] task provider.

Moreover, the author claims that the decisions falling in the upper left quadrant (e.g. medical AI applications) can be labelled as complex, given those require high cognitive demands concerning understanding and execution from the human [21]. High task complexity results in simplified judgement based on the concept of cut-off points, where it has been demonstrated that an option is excluded because of the individual rejection threshold [15]. This finding is in line with high arousal fostering simplified decision-making. Therefore, high arousal suppresses rational thoughts on AI performance (e.g. the fact that the AI provider outperforms the medical professional [105]), which in turn discourages the human to delegate the responsibility of task performance to an AI-based agent. Additionally, Campbell [21] noted that task complexity triggers subjects to recall already existing heuristics on how to approach the particular problem. In other words, the human is likely to refrain from AI use due to minor or even non-existent conscious touchpoints with AI in the past [51], leading to disregard of AI as a potential solution from the outset. Thus, the contradiction between arousal and valence in terms of motivation to employ AI (negative valence favouring AI and high arousal discouraging AI application) results in a conditional utilization of AI-based solutions. Possible conditions drawn from the literature are outlined in the following paragraph.

Longoni et al. [71] introduced the concept of “uniqueness neglect”, which refers to the person’s conviction that his or her distinctiveness can only be understood and considered by a human, not AI. Thus, perceived uniqueness (if it is a dominant characteristic of the human) is a factor likely to discourage the human to forego individual task completion by AI. The author classifies uniqueness into the category of person-related factors, next to technology- and context-related factors. Those factors are based on the existing research on AI and general technology aversion and constitute “the condition” in the upper left quadrant. Another example for a person-related factor is the need for control. For instance, humans have been found to be more willing to employ AI-based solutions for tasks where the human can even slightly impact the outcome [38]. However, need for control is a subjective variable in humans, reflecting aspects such as cultural heritage and personal experience [65]. Therefore, a person in the upper left quadrant, with a pronounced need for control,

will only be willing to apply AI-based task completion when the option to impact the task delegation process exists. Consequently, the factor “perceived disempowerment” can have an additional negative impact on readiness for task completion by AI [120].

An example of a context-related factor is the consequentiality of the task. The more consequences the human attributes to the particular task, the higher his or her perceived riskiness, and the less likely he or she is to rely on AI-based solutions [34], since the perceived sacrifice is too great [7]. The technology-related factor “competence” is strongly correlated with the exemplified context-related factors. Precisely, the perceived level of competence in the choice context drives the decision for or against an AI-based solution [9], [72]. Moreover, the factor “anthropomorphism” plays a significant role in the human judgement process, as anthropomorphized AI solutions are more accepted by humans [39], [90]. Thus, if the degree of humanization is a conditional factor for the respective person in the upper left quadrant, he or she is more likely to accept AI delegation in the case of an anthropomorphized AI solution. Overall, humans in the upper left quadrant decide on autonomy renouncement for a particular task based on the individualized impact of person-, technology-, and context-related factors. Nevertheless, given the arousing and unpleasant emotional circumstances of the decision situation at hand, the affective state of the human shapes the underlying evaluation process surrounding AI applicability.

In summary, the affective state of humans in the decision situation determines their initial motivation to forego execution in favour of AI delegation. The resulting impact of emotions on decision-making can be altered by customer-, technology-, or context-specific factors, whereby the impact of these factors differs based on the respective “emotional quadrant” the person is in.

6 Conclusion

AI has been shown to have the potential to significantly transform the human lifestyle [23]. This conceptual paper provides a taxonomy on human–AI interaction based on the theory of emotion-driven decision-making and concludes that the individual’s emotional state, a combination of valence and arousal, drives the motivation in favour of or against AI services. These insights contribute to the established research on the driving forces behind human decision-making and add to the understanding of human decision-making in the era of digitalization. For future human–AI interaction, it is necessary to understand the decisive factors concerning AI delegation and that is the main contribution of this paper. The author hopes that her discussion of human autonomy in the era of AI will inspire researchers on subsequent work in this dynamic research area.

References

1. Abraham, H., Lee, C., Brady, S., Fitzgerald, C., Mehler, B., Reimer, B., Coughlin, J. F.: Autonomous vehicles, trust, and driving alternatives: a survey of consumer preferences (2016)
2. Abraham, H., Reimer, B., Seppelt, B., Fitzgerald, C., Mehler, B., Coughlin, J. F.: Consumer interest in automation: Change over one year. In: Proceedings of the Transportation Research Board 97th Annual Meeting, 1–20 (2018)
3. Achar, C., So, J., Agrawal, N., Duhachek, A.: What we feel and why we buy: the influence of emotions on consumer decision-making. *Curr. Opin. Psychol.* **10**(1), 166–170 (2016)
4. Aggarwal, P., Mazumdar, T.: Decision delegation: a conceptualization and empirical investigation. *Psychol. Mark.* **25**(1), 71–93 (2008)
5. Alpert, S.R., Karat, J., Karat, C.M., Brodie, C., Vergo, J.G.: User attitudes regarding a user-adaptive eCommerce web site. *User Model. User-Adap. Inter.* **13**(4), 373–396 (2003)
6. Alvino, C., Basilico, J.: Learning a personalized homepage. *Netflix Technology Blog* (2015). <https://netflixtechblog.com/learning-a-personalized-homepage-aa8ec670359a>. Last accessed 21 Apr 2021
7. Ameen, N., Tarhini, A., Reppel, A., Anand, A.: Customer experiences in the age of artificial intelligence. *Comput. Human Behavior* **114**(1), 106548 (2021)
8. Andrade, E.B., Ariely, D.: The enduring impact of transient emotions on decision making. *Organ. Behav. Hum. Decis. Process.* **109**(1), 1–8 (2009)
9. André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., Van Boven, L., Weber, B., Yang, H.: Consumer choice and autonomy in the age of artificial intelligence and big data. *Cust. Needs Solut.* **5**(1–2), 28–37 (2018)
10. Aw, D. J.: When health care is horrifying: treating a patient with a fear of MDs. *National Post* (2012). <https://nationalpost.com/health/when-health-care-is-horrifying-treating-a-patient-with-a-fear-of-doctors>. Last accessed 29 Apr 2021
11. Azer, J., Alexander, M.J.: Conceptualizing negatively valenced influencing behavior: forms and triggers. *J. Serv. Manag.* **29**(3), 468–490 (2018)
12. Banovsky, M.: How Are You Keeping A Passion For Driving Alive? *Petrolicious* (2016). <https://petrolicious.com/articles/how-are-you-keeping-a-passion-for-driving-alive>. Last accessed 07 May 2021
13. Barrett, L.F.: Solving the emotion paradox: categorization and the experience of emotion. *Personal. Soc. Psychol. Rev.* **10**(1), 20–46 (2006)
14. Baum, S.D., Goertzel, B., Goertzel, T.G.: How long until human-level AI? results from an expert assessment. *Technol. Forecast. Soc. Chang.* **78**(1), 185–195 (2011)
15. Beach, L.R.: Broadening the definition of decision making: the role of prechoice screening of options. *Psychol. Sci.* **4**(4), 215–220 (1993)
16. Bechara, A.: The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain Cogn.* **55**(1), 30–40 (2004)
17. Bechara, A., Damasio, H., Damasio, A.R.: Emotion, decision making and the orbitofrontal cortex. *Cereb. Cortex* **10**(3), 295–307 (2000)
18. Beisswingert, B.M., Zhang, K., Goetz, T., Fischbacher, U.: Spillover effects of loss of control on risky decision-making. *PLoS ONE* **11**(3), 0150470 (2016)
19. Bosch, E., Ihme, K., Drewitz, U., Jipp, M., & Oehl, M.: Why drivers are frustrated: results from a diary study and focus groups. *Euro. Transp. Res. Rev.* **12**(1), (2020)
20. Bryk, W.: Artificial superintelligence: the coming revolution. *Harvard Sci. Rev.* (2015). <https://harvardsciencereview.org/2015/12/04/artificial-superintelligence-the-coming-revolution-2/>. Last accessed 11 Jun 2021
21. Campbell, D.J.: Task complexity: a review and analysis. *Acad. Manag. Rev.* **13**(1), 40–52 (1988)
22. Canziani, B., MacSween, S.: Consumer acceptance of voice-activated smart home devices for product information seeking and online ordering. *Comput. Human Behavior* **119**(1), 106714 (2021)

23. Carter, S., Nielsen, M.: Using artificial intelligence to augment human intelligence. *Distill* **2**(12), 9 (2017)
24. Castelo, N.: *Blurring the line between human and machine: marketing artificial intelligence*. ProQuest Dissertations Publishing, New York (2019)
25. Castelo, N., Bos, M.W., Lehmann, D.R.: Task-dependent algorithm aversion. *J. Mark. Res.* **56**(5), 809–825 (2019)
26. Chan, E.S.W., Okumus, F., Chan, W.: The applications of environmental technologies in hotels. *J. Hosp. Market. Manag.* **26**(1), 23–47 (2017)
27. Christie's: Is artificial intelligence set to become art's next medium? (2018) <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>. Last accessed 15 Mar 2021
28. Clarke, K., Belk, R.W.: The effects of product involvement and task definition on anticipated consumer effort. *Adv. Consum. Res.* **6**(1), 313–318 (1979)
29. Clarke, L.: Domino's Pizza bets on the future of voice ordering with Twilio. *Channel Asia* (2018). <https://www.channelasia.tech/article/648623/domino-pizza-bets-future-voice-ordering-twilio/?fp=2&fpid=1>. Last accessed 11 Jun 2021
30. Constantinou, E., Bogaerts, K., Van Diest, I., Van den Bergh, O.: Inducing symptoms in high symptom reporters via emotional pictures: The interactive effects of valence and arousal. *J. Psychosom. Res.* **74**(3), 191–196 (2013)
31. Conti, S., Gourvenec, B., Billot, R.: Impact of artificial intelligence in customer service: state of the art. In: *Proceedings of the Global Marketing Conference at Seoul*, 807–823 (2020)
32. Costa, H., Macedo, L.: Emotion-based recommender system for overcoming the problem of information overload. In J. M. Corchado, J. Bajo, J. Kozlak, P. Pawlewski, J. M. Molina, V. Julian, R. Azambuja Silveira, R. Unland, S. Giroux (eds.), *Highlights on Practical Application of Agents and Multi-agent Systems*, 178–189 (2013)
33. Danise, A.: Women say no thanks to driverless cars, survey finds; Men say tell me more. *Nerd-Wallet* (2015). <https://www.nerdwallet.com/blog/insurance/auto/survey-consumer-fears-self-driving-cars/>. Last accessed 22 Apr 2021
34. Davenport, T., Guha, A., Grewal, D., Bressgott, T.: How artificial intelligence will change the future of marketing. *J. Acad. Mark. Sci.* **48**(1), 24–42 (2020)
35. De Bellis, E., Johar, G.V.: Autonomous shopping systems: Identifying and overcoming barriers to consumer adoption. *J. Retail.* **96**(1), 74–87 (2020)
36. Diab, D.L., Pui, S.Y., Yankelevich, M., Highhouse, S.: Lay perceptions of selection decision aids in US and non-US samples. *Int. J. Sel. Assess.* **19**(2), 209–216 (2011)
37. Dietvorst, B.J., Simmons, J.P., Massey, C.: Algorithm aversion: people erroneously avoid algorithms after seeing them err. *J. Exp. Psychol.* **144**(1), 114–126 (2015)
38. Dietvorst, B.J., Simmons, J.P., Massey, C.: Overcoming algorithm aversion: people will use imperfect algorithms if they can (even slightly) modify them. *Manage. Sci.* **64**(3), 983–1476 (2018)
39. DiSalvo, C., Gemperle, F., Forlizzi, J., Montgomery, E.: The hug: an exploration of robotic form for intimate communication. In: *Proceedings of the 12th International Workshop on Robot and Human Interactive Communication*, 403–408 (2003)
40. Eastwood, J., Snook, B., Luther, K.: What people want from their professionals: attitudes toward decision-making strategies. *J. Behav. Decis. Mak.* **25**(5), 458–468 (2012)
41. Gendron, M., Feldman Barrett, L.: Reconstructing the past: a century of ideas about emotion in psychology. *Emotion Rev.* **1**(4), 316–339 (2009)
42. Grace, K., Salvatier, J., Dafoe, A., Zhang, B., Evans, O.: Viewpoint: when will AI exceed human performance? evidence from AI experts. *J. Artif. Intell. Res.* **62**, 729–754 (2018)
43. Guo, X., Shen, Z., Zhang, Y., Wu, T.: Review on the application of artificial intelligence in smart homes. *Smart Cities* **2**(3), 402–420 (2019)
44. Gursoy, D., Chi, O.H., Lu, L., Nunkoo, R.: Consumers acceptance of artificially intelligent (AI) device use in service delivery. *Int. J. Inf. Manage.* **49**(1), 157–169 (2019)
45. Haboucha, C.J., Ishaq, R., Shifan, Y.: User preferences regarding autonomous vehicles. *Transp. Res. Part C: Emerg. Technol.* **78**(1), 37–49 (2017)

46. Haenlein, M., Kaplan, A.: A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. *Calif. Manage. Rev.* **61**(4), 5–14 (2019)
47. Häubl, G., Trifts, V.: Consumer decision making in online shopping environments: the effects of interactive decision aids. *Mark. Sci.* **19**(1), 4–21 (2000)
48. Hohenberger, C., Spörrle, M., Welpe, I.M.: How and why do men and women differ in their willingness to use automated cars? the influence of emotions across different age groups. *Transp. Res. Part A: Policy Pract.* **94**(1), 374–385 (2016)
49. Hollander, M.A.G., Greene, M.G.: A conceptual framework for understanding iatrophobia. *Patient Educ. Couns.* **102**(11), 2091–2096 (2019)
50. Howard, D., Dai, D.: Public perceptions of self-driving cars. *Transp. Res. Board 93rd Ann. Meet.* (2014)
51. Hyken, S.: Half of people who encounter artificial intelligence don't even realize it. *Forbes* (2017). <https://www.forbes.com/sites/shephyken/2017/06/10/half-of-people-who-encounter-artificial-intelligence-dont-even-realize-it/?sh=340a96bc745f>. Last accessed 19 Mar 2021
52. Isen, A.M.: Positive affect, cognitive processes, and social behavior. *Adv. Exp. Soc. Psychol.* **20**, 203–253 (1987)
53. Jussupow, E., Benbasat, I., Heinzl, A.: Why are we averse towards algorithms? a comprehensive literature review on algorithm aversion. In: *Proceedings of the 28th European Conference on Information Systems* (2020)
54. Kahneman, D., Tversky, A.: Choices, values, and frames. *Am. Psychol.* **39**(4), 341–350 (2013)
55. Kapp, B. S., Whalen, P. J., Supple, W. F., & Pascoe, J. P.: Amygdaloid contributions to conditioned arousal and sensory information processing. *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction.* 229–254 (1992)
56. König, M., Neumayr, L.: Users' resistance towards radical innovations: the case of the self-driving car. *Transport. Res. F: Traffic Psychol. Behav.* **44**(1), 42–52 (2017)
57. Kowalczyk, P.: Consumer acceptance of smart speakers: a mixed methods approach. *J. Res. Interact. Mark.* **12**(4), 418–431 (2018)
58. Kowalski, J., Biele, C., Krzysztofek, K.: Smart home technology as a creator of a super-empowered user. In Karwowski W, Ahram T. (eds.), *Intelligent Human Systems Integration 2019*, pp. 175–180. Springer (2019)
59. Kranz, J., Gallenkamp, J., Picot, A.: Exploring the role of control—smart meter acceptance of residential consumers. In: *Proceedings of the 16th Americas Conference on Information Systems* (2010)
60. Kunz, W.H., Heinonen, K., Lemmink, J.G.A.M.: Future service technologies: Is service research on track with business reality? *J. Serv. Mark.* **33**(4), 479–487 (2019)
61. Kyriakidis, M., Happee, R., De Winter, J.C.F.: Public opinion on automated driving: results of an international questionnaire among 5000 respondents. *Transport. Res. F: Traffic Psychol. Behav.* **32**(1), 127–140 (2015)
62. Lang, A., Chung, Y., Lee, S., Zhao, X.: It's the product: do risky products compel attention and elicit arousal in media users? *Health Commun.* **17**(3), 283–300 (2005)
63. Lee, Y., Kozar, K., Larsen, K.R.: The technology acceptance model: past, present, and future. *Commun. Assoc. Inf. Syst.* **12**(1), 50 (2003)
64. Leino, J., Riihjä, K. J.: Case Amazon: Ratings and reviews as part of a recommendations. In: *RecSys'07 Proceedings of the 2007 ACM Conference on Recommender Systems*, 137–140 (2007)
65. Leotti, L.A., Iyengar, S.S., Ochsner, K.N.: Born to choose: the origins and value of the need for control. *Trends Cogn. Sci.* **14**(10), 457–463 (2010)
66. Leung, E., Paolacci, G., Puntoni, S.: Man versus machine: resisting automation in identity-based consumer behavior. *J. Mark. Res.* **55**(6), 818–831 (2018)
67. Lewinsohn, S., Mano, H.: Multi-attribute choice and affect: the influence of naturally occurring and manipulated moods on choice processes. *J. Behav. Decis. Mak.* **6**(1), 33–51 (1993)
68. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Comput.* **7**(1), 76–80 (2003)

69. Lindquist, K.A., Barrett, L.F.: Constructing emotion: the experience of fear as a conceptual act. *Psychol. Sci.* **19**(9), 898–903 (2008)
70. Linville, P.W., Jones, E.E.: Polarized appraisals of out-group members. *J. Pers. Soc. Psychol.* **38**(5), 689–703 (1980)
71. Longoni, C., Bonezzi, A., Morewedge, C.: Resistance to medical artificial intelligence. *J. Consumer Res.* **46**(4), 629–650 (2019)
72. Longoni, C., Cian, L.: Artificial Intelligence in utilitarian vs. hedonic contexts: The “Word-of-Machine” effect. *J. Market.* (2020)
73. Lu, L., Cai, R., Gursoy, D.: Developing and validating a service robot integration willingness scale. *Int. J. Hosp. Manag.* **80**(1), 36–51 (2019)
74. Mano, H.: Emotional states and decision making. *Adv. Consum. Res.* **17**(1), 577–584 (1990)
75. Mano, H.: Judgments under distress: Assessing the role of unpleasantness and arousal in judgment formation. *Organ. Behav. Hum. Decis. Process.* **52**(2), 216–245 (1992)
76. Mano, H.: Risk-taking, framing effects, and affect. *Organ. Behav. Hum. Decis. Process.* **51**(1), 38–58 (1994)
77. Mano, H.: Affect and persuasion: The influence of pleasantness and arousal on attitude formation and message elaboration. *Psychol. Mark.* **14**(4), 315–335 (1997)
78. Mehrabian, A., Russell, J.A.: An approach to environmental psychology. *Electronics Power* **18**(5), 167–170 (1974)
79. Meng, Y., Zhang, W., Zhu, H., Shen, X.S.: Securing consumer IoT in the smart home: architecture, challenges, and countermeasures. *IEEE Wirel. Commun.* **25**(6), 53–59 (2018)
80. Minh, V.T., Khanna, R.: Application of artificial intelligence in smart kitchen. *Int. J. Innov. Technol. Interdisciplin. Sci.* **1**(1), 1–8 (2018)
81. Mills, R.C.: A new understanding of self: the role of affect, state of mind, self-understanding, and intrinsic motivation. *Journal of Experimental Education* **60**(1), 67–81 (1991)
82. Mitchell, A.: Involvement: a potentially important mediator of consumer behavior. *Adv. Consumer Res.* **6**(1), 191–196 (1979). <http://acrwebsite.org/volumes/5703/volumes/v06/NA-06>. Last accessed 14 May 2021
83. Mitchell, D.: The nexus between decision making and emotion regulation: a review of convergent neurocognitive substrates. *Behav. Brain Res.* **217**(1), 215–231 (2011)
84. Mizgajski, J., Morzy, M.: Affective recommender systems in online news industry: how emotions influence reading choices. *User Model. User-Adap. Inter.* **29**(2), 345–379 (2019)
85. Müller, V.C., Bostrom, N.: Future progress in artificial intelligence: a survey of expert opinion. *Fundamental Issues Artific. Intell.*, 555–572 (2016)
86. Nass, C., Jonsson, I. M., Harris, H., Reaves, B., Endo, J., Brave, S., Takayama, L.: Improving automotive safety by pairing driver emotion and car voice emotion. In: *Proceedings of the Conference on Human Factors in Computing Systems, 1973–1976* (2005)
87. Nawijn, J., Mitas, O., Lin, Y., Kerstetter, D.: How do we feel on vacation? a closer look at how emotions change over the course of a trip. *J. Travel Res.* **52**(2), 265–274 (2012)
88. Nielsen: (Smart) Speaking my language: despite their vast capabilities, smart speakers are all about the music (2018). <https://www.nielsen.com/us/en/insights/article/2018/smart-speaking-my-language-despite-their-vast-capabilities-smart-speakers-all-about-the-music/>. Last accessed 05 May 2021
89. Niven, K.: Affect. In Gellman, M. D., Turner, J. R. (eds.), *Encyclopedia of Behavioral Medicine*, pp. 49–50. Springer (2013)
90. Nowak, K. L., Biocca, F.: The effect of the agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators Virtual Environ.* **12**(5), 481–494 (2003)
91. Ostrom, A. L., Fotheringham, D., Bitner, M. J.: Customer acceptance of AI in service encounters: understanding antecedents and consequences. In Maglio, P. P., Kieliszewski, C. A., Spohrer, J. C., Lyons, K., Patrício, L., Sawatani, Y. (eds.), *Handbook of Service Science*, pp. 77–103. Springer (2019)
92. Ovide, S.: Driverless cars go humble to get real. *The New York Times* (2020). <https://www.nytimes.com/2020/10/26/technology/driverless-cars.html>. Last accessed 21 Apr 2021

93. Pajkovic, N.: Algorithms and taste-making: exposing the Netflix recommender system's operational logics. *Convergence* (2021)
94. Palmeira, M., Spassova, G.: Consumer reactions to professionals who use decision aids. *Eur. J. Mark.* **49**(3/4), 302–326 (2015)
95. Park, E., Kim, S., Kim, Y.S., Kwon, S.J.: Smart home services as the next mainstream of the ICT industry: determinants of the adoption of smart home services. *Univ. Access Inf. Soc.* **17**(1), 175–190 (2018)
96. Park, K., Kwak, C., Lee, J., Ahn, J.H.: The effect of platform characteristics on the adoption of smart speakers: empirical evidence in South Korea. *Telematics Inform.* **35**(8), 2118–2132 (2018)
97. Paulus, M.P., Yu, A.J.: Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* **16**(9), 476–483 (2012)
98. Payre, W., Cestac, J., Delhomme, P.: Intention to use a fully automated car: attitudes and a priori acceptability. *Transp. Res. Part F: Traffic Psychol. Behav.*, **27**(B), 252–263 (2014)
99. Peters, E., Västfjäll, D., Gärling, T., Slovic, P.: Affect and decision making: a “hot” topic. *J. Behav. Decis. Mak.* **19**(2), 79–85 (2006)
100. Petty, R. E., Cacioppo, J. T.: The elaboration likelihood model of persuasion. In Petty, R. E., Cacioppo, J. T. (Eds.), *Communication and Persuasion*, pp. 1–24. Springer (1986).
101. Pfister, H.R., Böhm, G.: The multiplicity of emotions: a framework of emotional functions in decision making. *Judgm. Decis. Mak.* **3**(1), 5–17 (2008)
102. Phelps, E.A., LeDoux, J.E.: Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**(2), 175–187 (2005)
103. Porter, M. E., Heppelmann, J. E.: How smart, connected products are transforming competition. *Harvard Bus. Rev.* (2014). <https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition>. Last accessed 11 Jun 2021
104. Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* **17**(3), 715–734 (2005)
105. Promberger, M., Baron, J.: Do patients trust computers? *J. Behav. Decis. Mak.* **19**(5), 455–468 (2006)
106. Raue, M., D’Ambrosio, L.A., Ward, C., Lee, C., Jacquillat, C., Coughlin, J.F.: The influence of feelings while driving regular cars on the perception and acceptance of self-driving cars. *Risk Anal.* **39**(2), 358–374 (2019)
107. Reed, A., Forehand, M.R., Puntoni, S., Warlop, L.: Identity-based consumer behavior. *Int. J. Res. Mark.* **29**(4), 310–321 (2012)
108. Reeve, J.: The interest-enjoyment distinction in intrinsic motivation. *Motiv. Emot.* **13**(2), 83–103 (1989)
109. Reisenzein, R.: Pleasure-arousal theory and the intensity of emotions. *J. Pers. Soc. Psychol.* **67**(3), 525–539 (1994)
110. Rijdsdijk, S.A., Hultink, E.J.: “Honey, have you seen our hamster?” consumer evaluations of autonomous domestic products. *J. Prod. Innov. Manag.* **20**(3), 204–216 (2003)
111. Rijdsdijk, S.A., Hultink, E.J.: How today’s consumers perceive tomorrow’s smart products. *J. Prod. Innov. Manag.* **26**(1), 24–42 (2009)
112. Risteska Stojkoska, B.L., Trivodaliev, K.V.: A review of internet of things for smart home: challenges and solutions. *J. Clean. Prod.* **140**, 1454–1464 (2017)
113. Russell, J.A.: A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**(6), 1161–1178 (1980)
114. Russell, J.A.: Core affect and the psychological construction of emotion. *Psychol. Rev.* **110**(1), 145–172 (2003)
115. Schafer, J. B., Konstan, J., & Riedl, J.: Recommender systems in e-commerce. In: *Proceedings of the 1st ACM Conference on Electronic Commerce*, 158–166 (1999).
116. Schmelzer, R.: Should we be afraid of AI? *Forbes* (2019). <https://www.forbes.com/sites/cognitiveworld/2019/10/31/should-we-be-afraid-of-ai/?sh=569bc8b4331d>. Last accessed 16 Mar 2021

117. Schoettle, B., Sivak, M.: A survey of public opinion about connected vehicles in the U.S., the U.K., and Australia. In: Proceedings of the 2014 International Conference on Connected Vehicles and Expo, 687–692 (2014)
118. Schwartz, E. H.: Voice assistants see uptick in daily use during the pandemic: Report. Voicebot (2020). <https://voicebot.ai/2020/05/01/voice-assistants-see-uptick-in-daily-use-during-the-pandemic-report/>. Last accessed 21 Apr 2021
119. Schwarz, N., Clore, G.L.: Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* **45**(3), 513–523 (1983)
120. Schweitzer, F., Van den Hende, E.A.: To be or not to be in thrall to the march of smart products. *Psychol. Mark.* **33**(10), 830–842 (2016)
121. Seymour, B., Dolan, R.: Emotion, decision making, and the amygdala. *Neuron* **58**(5), 662–671 (2008)
122. Shaffer, V.A., Probst, C.A., Merkle, E.C., Arkes, H.R., Medow, M.A.: Why do patients derogate physicians who use a computer-based diagnostic support system? *Med. Decis. Making* **33**(1), 108–118 (2013)
123. Shankar, V.: How Artificial Intelligence (AI) is reshaping retailing. *J. Retail.* **94**(4), vi–xi (2018)
124. Sharar, S.R., Alamdari, A., Hoffer, C., Hoffman, H.G., Jensen, M.P., Patterson, D.R.: Circumplex model of affect: a measure of pleasure and arousal during virtual reality distraction analgesia. *Games Health J.* **5**(3), 197–202 (2016)
125. Shetty, D.: Women Across 17 Countries Report Increased Stress, Household Work Due To Covid-19. *Forbes* (2021). <https://www.forbes.com/sites/dishashetty/2021/01/29/women-across-17-countries-report-increased-stress-household-work-due-to-covid-19/?sh=4ece1b852f0a>. Last accessed 03 Jun 2021
126. So, J., Achar, C., Han, D.H., Agrawal, N., Duhachek, A., Maheswaran, D.: The psychology of appraisal: specific emotions and decision-making. *J. Consum. Psychol.* **25**(3), 359–371 (2015)
127. Sohn, K., Kwon, O.: Technology acceptance theories and factors influencing artificial intelligence-based intelligent products. *Telematics Inform.* **47**(December), 1–14 (2020)
128. Stein, M.B., Simmons, A.N., Feinstein, J.S., Paulus, M.P.: Increased amygdala and insula activation during emotion processing in anxiety-prone subjects. *Am. J. Psychiatry* **164**(2), 318–327 (2007)
129. Svebak, S., Stoyva, J.: High arousal can be pleasant and exciting: the theory of psychological reversals. *Biofeedback Self Regul.* **6**(1), 443 (1981)
130. Tkalčić, M., Burnik, U., Košir, A.: Using affective parameters in a content-based recommender system for images. *User Model. User-Adap. Inter.* **20**(4), 279–311 (2010)
131. Tseng, A., Bansal, R., Liu, J., Gerber, A.J., Goh, S., Posner, J., Colibazzi, T., Algermissen, M., Chiang, I.-C., Russel, J.A., Peterson, B.S.: Using the circumplex model of affect to study valence and arousal ratings of emotional faces by children and adults with autism spectrum disorders. *J. Autism Dev. Disord.* **44**(6), 1332–1346 (2014)
132. Van’t Wout, M., Kahn, R. S., Sanfey, A. G., Aleman, A.: Affective state and decision-making in the Ultimatum Game. *Experiment. Brain Res.*, **169**(4), 564–568 (2006)
133. Verganti, R., Vendraminelli, L., Iansiti, M.: Innovation and design in the age of Artificial Intelligence. *J. Prod. Innov. Manag.* **37**(3), 212–227 (2020)
134. Wang, X., McGill, T.J., Klobas, J.E.: I want it anyway: consumer perceptions of smart home devices. *J. Comput. Inf. Syst.* **60**(5), 437–447 (2020)
135. Watson, D., Tellegen, A.: Toward a consensual structure of mood. *Psychol. Bull.* **98**(2), 219–235 (1985)
136. Wei, K., Huang, J., Fu, S.: A survey of E-commerce recommender systems. In: Proceedings of the 2007 International Conference on Service Systems and Service Management (2007)
137. Williams, R.: Domino’s lets AI assistant “Dom” handle incoming phone orders. *Marketing Dive* (2018). <https://www.marketingdive.com/news/dominos-lets-ai-assistant-dom-handle-incoming-phone-orders/522163/>. Last accessed 17 Mar 2021

138. Yang, H., Lee, H., Zo, H.: User acceptance of smart home services: an extension of the theory of planned behavior. *Ind. Manag. Data Syst.* **117**(1), 68–89 (2017)
139. Zajonc, R.B.: The interaction of affect and cognition. *Approach. Emotion* **239**(1), 239–246 (1984)
140. Zeelenberg, M., Nelissen, R.M.A., Breugelmans, S.M., Pieters, R.: On emotion specificity in decision making: Why feeling is for doing. *Judgm. Decis. Mak.* **3**(1), 18–27 (2008)

Roadmapping Collaborative Exploitation and Marketing of an AI-Based Knowledge Platform



Andrzej M. J. Skulimowski 

Abstract Accelerated development of open-access, web-based information systems is an important driver of social and economic change in knowledge-based economies. Specifically, learning platforms and knowledge repositories employing modern AI tools (AILPs) strongly influence human resources management in academic institutions and corporations. AI-based information systems require efficient alignment of business models to the deployment of digital technologies, where platform marketing plays a crucial role. This is why building an appropriate marketing and exploitation strategy is fundamental in achieving the desired economic impacts of AILP operation and exploitation, and for financial decisions concerning further AILP development. This paper presents the methods and final outcomes of collaborative roadmapping as part of an exploitation and marketing strategy and user community building plan for an AILP implemented within a recent Horizon 2020 project. Platform marketing activities have a hierarchical structure. They touch upon the platform as a whole, the platform as a learning tool for its users, as well as may aim at licensing some of its stand-alone software components. The overall strategy has been composed of individual and joint activities. The strategy-building methodology consists of generating alternative action plans, evaluated by multiple criteria. The compromise action plan can be adaptively modified according to user preferences and stakeholder needs. Finally, we will present how the roadmapping diagram was built and assessed to yield a realistic strategy.

Keywords AI-based learning platforms · Collaboration roadmapping · Marketing · Information systems · Exploitation strategy · Economic impacts

A. M. J. Skulimowski (✉)

Decision Sciences Laboratory, Chair of Automatic Control and Robotics, AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: ams@agh.edu.pl

International Centre for Decision Sciences and Forecasting, Progress & Business Foundation, ul. Lea 12B, 30-048 Kraków, Poland

1 Introduction

The rise of artificial intelligence (AI) methods, technologies, as well as their applications to digital learning platforms (AILPs) and knowledge repositories, poses a challenge to software developers in informing potentially interested customers about new products and their features. On the other hand, web system owners and operators are obliged to inform current and potential users and attract them to communities of practice around their platform. Both types of marketing are particularly relevant for novel and prototype systems, developed within research and development (R&D) projects by interdisciplinary consortia. System alignment to current information technologies (IT), marketing, user community building and exploitation requires thorough strategy building. Such a strategy will be further referred to as an exploitation strategy.

This paper presents the methodology of joint marketing and exploitation strategy building and its collaborative implementation by a platform developer consortium. Collaborative roadmapping turned out to be a fundamental component of strategy building for an AILP, as evidenced within a recent EU Horizon 2020 project (acronym MOVING, www.moving-project.eu, [13]). Marketing activities resulting from this strategy have a hierarchical structure with two principal levels:

Level 1. The marketing of a platform as a whole, to find new customers for licensing and new stakeholders for the existing platform.

Level 2. User community building for the platform, considered as a learning tool.

There is also an ancillary level, which can be specified as:

Level 3. Marketing of principal software components and licensing of independent software modules (ISMs) and algorithms.

Finally, we can define the meta level of marketing and promotional activities, which partly overlaps with levels 1 and 3 and the dissemination of research results.

Level 0. Activities geared towards finding funds for the continual development of the platform and related research.

A conceptual model of the overall platform marketing is shown in Fig. 1.

The latter meta level of marketing activities may involve seeking venture capital as well as additional research grants. In the case of the platform developed during the project MOVING, grants can be sought predominantly in Horizon Europe as well as in other available EU research and innovation fostering programmes such as Digital Europe. Most of the remaining marketing, promotional and research dissemination may be considered as supporting activities for the achievement of this goal. These activities include user preference elicitation, publications in professional and scholarly journals, conference and exhibition presentations, campaigns in social media, etc.

As noted above, the strategy-building methodology has been applied to the exploitation, user community-building and marketing action planning of an AILP, developed within the aforementioned Horizon 2020 research project. This repository

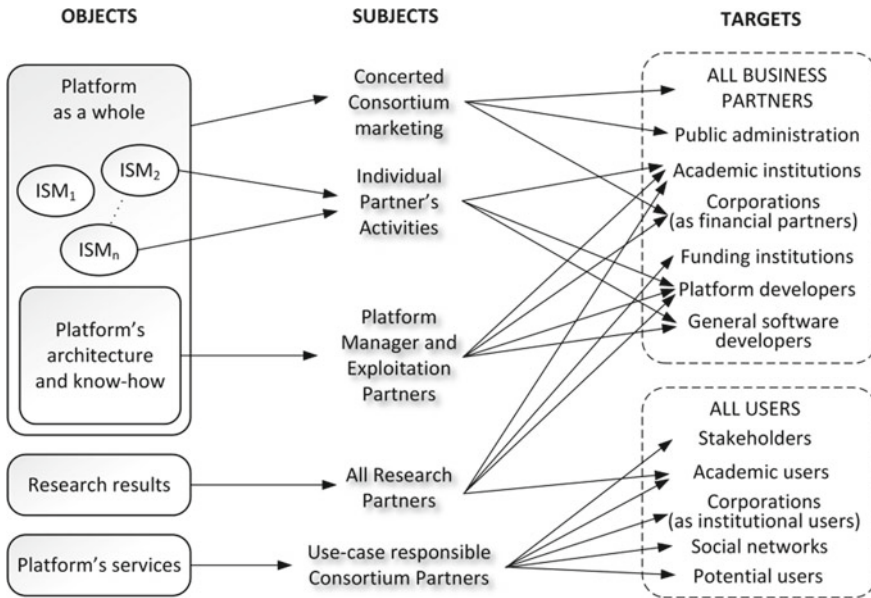


Fig. 1 A conceptual scheme of learning platform marketing. ISM_{*i*} denotes the *i*th individual software module, offered as a separate product

stores online courses, scholarly papers, dynamically updated economic information, financial and other data. It can be used for the development of human resources, research and supporting business processes. We will refer to this AILP as ‘the platform’.

The overall platform marketing strategy is composed of individual as well as concerted activities presented in subsequent sections. In Sect. 2, we provide a brief overview of related research and the basics of background methodology. Section 3 presents how the roadmapping diagram (RD) was built and assessed to yield a realistic cooperation strategy for the developer consortium. This section also discusses the related roadmapping analytics. Final conclusions are presented in Sect. 4.

2 Related Research

Very few descriptions of AILP strategies or strategy building approaches exist. Those available refer mostly to digital libraries [6] and e-learning course repositories [2]. This lack of available tools necessitated the development of methodological foundations for AILP-oriented strategic planning, to build new ICT-based tools and collaborative approaches, and apply them to satisfy project goals within the context of EU research and innovation policies. A methodological outline of strategy building was presented in [8]. In [9], a novel expert knowledge elicitation and processing

tool that builds on the Delphi survey methodology to construct an AILP exploitation strategy and estimate its economic impact was presented. The background of expert extrapolation Delphi is described in [7]. Its hybrid multiround—real-time implementation can be found at www.forgnosis.eu. Another relevant tool is collaborative roadmapping, which has been adopted from classical technological roadmapping techniques [14], cf. also [4] for the history of the roadmapping development and [1] for the roadmapping variant integrating marketing and strategic planning. Information provided by experts as responses to a Delphi survey, together with stakeholder and developer contributions to defining the roadmap objects, yields a set of plausible scenarios, which are then assessed by the roadmapping facilitator [12], denoted as RdF in RD.

Further clues regarding AILP marketing can be provided by the general theory of information platforms, cf. e.g. [5]. However, this theory is biased by the analysis of large e-commerce social media, information exchange platforms and search engines [15]. Therefore, it should be used with precaution when applied to learning sites.

Recently, in [10] and [11], issues related to the assessment and optimization of the social impact of AILPs were analysed. These papers define the economic, business, market and business-oriented research environments of knowledge repositories.

The preference elicitation of potential platform users is another issue of fundamental importance. User preferences concerning human-platform communication and availability of content and training support tools influence the platform choice [3] and may drive developers to improve the design of graphical user interfaces (GUI) and create various AI tools to facilitate learning.

3 Collaboration-Oriented Roadmapping

From the AILP strategy building scheme shown in [9], it follows that the collaborative roadmapping exercise is the final stage of strategy building. According to the general roadmapping methodology, roadmap design is preceded by the definition of RD layers, layer objects and relations between them, as well as by data acquisition to quantify the diagram. This process uses a multivariate vision of the future in the form of scenarios derived from the Delphi survey and expert panels, from bibliometric, patentometric and webometric data, macroeconomic statistics, as well as other economic data sources. Publicly available prospective studies may be used to establish relations between different future factors [8, 9]. They serve, in turn, to identify a set of feasible actions out of a catalogue of exploitation and user community-building activities.

3.1 Preference Elicitation and Collaborative Roadmapping Diagram Building

A sample roadmapping diagram derived from a collaborative decision support process in a Horizon 2020 project [13] is shown in Fig. 2. Horizontal bars are associated with R&D and commercialization activities planned by the project partners P_i.

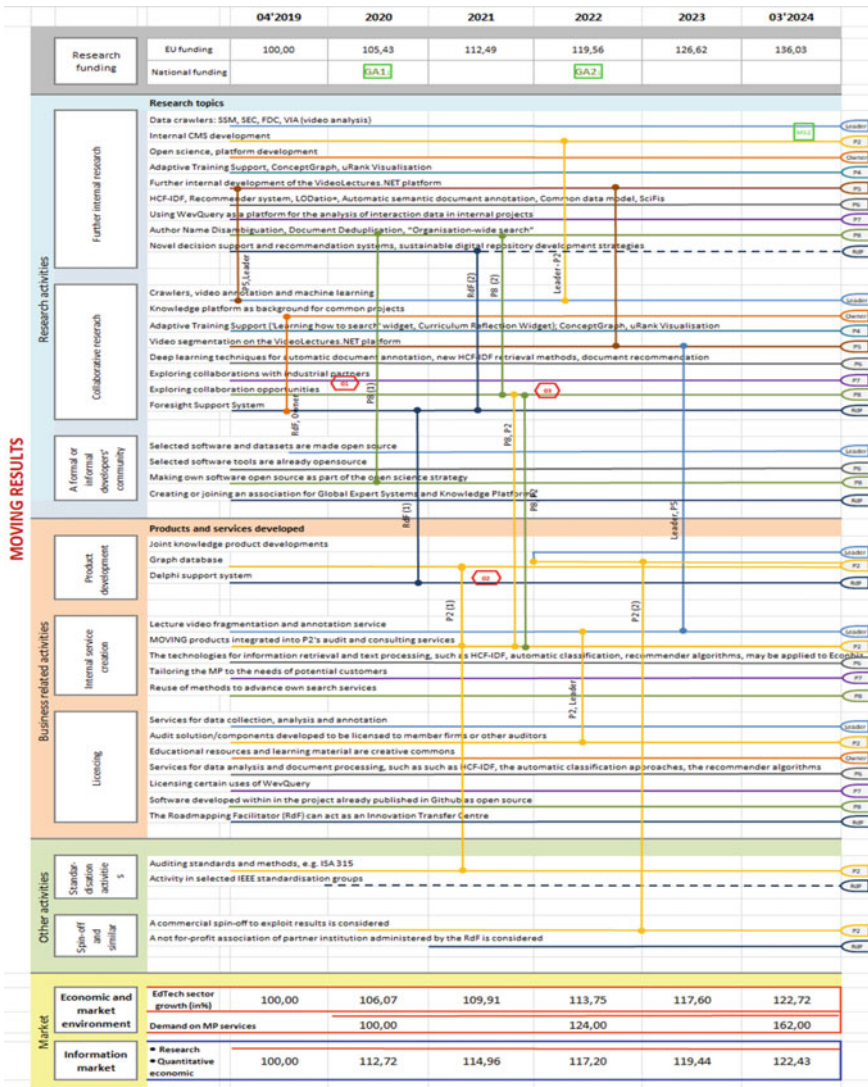


Fig. 2 A sample AILP exploitation roadmap. Horizontal bars denote individual partner tasks

Eight types of activities are grouped in three main layers. The other layers describe the market (bottom) and research funding opportunities (top). Vertical bars denote triggering conditions and other relations between partner activities, such as forming joint ventures. Red diamonds correspond to real expansion options, while the green rectangles represent expected research grants. Activity layers are determined by the organizations involved in the exercise during the periodic facilitated roadmapping sessions. The data describing the market and funding layers come from prospective studies, market research, official statistics and R&D policy statements. Activities and tasks which are predetermined at the beginning of the exercise and cannot be changed, or those which can only be changed by other (external) decision-makers, are termed *passive*. On the contrary, tasks which may be changed, abandoned or re-started are termed *active*.

By its construction, the exploitation roadmap cannot only consist of passive tasks. A passive task, if included, must be accompanied by a 'yield' element, i.e. an associated activity bringing a direct benefit to its owner. For example, '*exploring a possibility*' is passive, but '*finding a possibility*' is active when this action is completed. This is why '*exploring a possibility*', if depicted in the diagram, should be linked vertically to some other tasks at the expected time of '*finding the possibility*' and making use of it. The latter opportunity can be modelled in the diagram as a real option (in this case, a project extension option). It can be driven by the cumulated 'exploration' effort, estimated as the average monthly fixed costs of personnel employed, full- or part-time and the corresponding overheads. '*Exploring opportunities*' can also imply a growing excellence in technology, services or products offered, causing a growth in their value. The role of an analyst is to facilitate the roadmapping exercise by discovering real options, proposing and redefining vertical relations within an interactive process, as well as performing calculations to derive optimal action plans and scenarios.

3.2 The Role of Real Options in the Roadmapping Process

Many technological companies use real options to value IT and AI innovations and other intangibles, while the proper definition of underlying parameters and estimation of other option parameters may constitute a challenge. Nevertheless, one can utilize the cost valuation method, which yields the cumulative development costs with capital expenses and discounted future development as the technology value. If the fixed costs are known, which is often the case, it is possible to sum up both values. An estimated value of the technology may then be included as an underlying parameter in valuing the project, extended with academia-industry cooperation. Due to high uncertainty about the technology value before the project end, a rigorous real option valuation cannot usually be performed when starting the roadmapping exercise.

Rights gained during the repository operation as well as liabilities, can be, respectively, modelled by long or short real option positions. The iterative dependence of

future investment opportunities on previous outcomes will be modelled by nested real options and embedded into an anticipatory network (AN, cf. [10]) that makes it possible to model the expected consequences of admitting a proposed operation strategy. Real option modelling includes switching, abandonment, and continuation options, thus yielding a realistic model of a knowledge repository. Observe that the above roadmap indicates the relevance of real options and a need for their later evaluation with growing accuracy. Similar real options, as in case of the above expansion option, may model the availability of licensing opportunities, while continuation options may relate to the possibility of getting additional research funding.

The diversified scope of planned marketing actions contributes to a higher degree of complexity in strategy planning optimization. This was the reason for choosing a real option-oriented extension of the roadmapping methodology as a framework for the decision support to solve the above platform exploitation planning problem. As far as financial criteria are concerned, the real options turned out to be a natural and useful tool for describing the relations between different deployment variants of the platform components and services, as well as of the platform as a whole, represented as objects of the new product layer in the RD. The activity deployment plans correspond to the retained best-compromise planning scenarios and associated optimal financial yields. Moreover, the financial valuation of the operation plans can easily be combined with a prior SWOTC (SWOT analysis with Challenges as an additional dimension, cf. [9]) assessment of roadmapping objects.

3.3 Applying the Roadmapping Diagram to Marketing Action Planning

After taking into account all the information about the activities depicted in the above RD, we can derive priority activities to be performed in the first order of importance. The activity ranking may be regarded as a conclusion of the individual exploitation plans shown in the RD. The information to be analysed features the precedence and other relations between diagram objects, scope of marketing and other exploitation activities proposed in the consortium partners' individual plans and information about the social and economic environment, retrieved from other sources. To establish the precedence and priority, suitable methods of marketing campaign optimality analysis, often used in strategic planning, such as multivariate PERT diagrams, could be applied. Selected priority joint exploitation activities, derived from the above roadmapping diagram (Fig. 2) for the platform development project, which serves here as an illustration of our methodology [13], are shown in Table 1.

The collaboration strength, expressed as the number of cooperating partners for a joint activity, was used as a score. It was weighted with each activity duration and value.

The information needed to perform the roadmapping process and build the RD has been mostly provided by the Delphi survey [7] and by the consortium partners

Table 1 Priority joint exploitation activities of the AILP developers, derived from the collaborative roadmapping and other forward-looking activities

No.	Joint exploitation activities with the highest expected yield	Recommended to
1	The AILP can be a playground for testing new services and AI-based functionalities. Novel recommendation systems, decision pilots, creativity support systems and content-based multimedia processing may enhance the platform for the benefit of its users, while simultaneously providing an opportunity to test novel services with qualified users. These users will be attracted by the quality of services offered and promoted via social networks and media	All AILP developers, in particular, AILP manager (owner) P ₃ , project leader P ₁ , RdF—P ₉
2	Use the AILP as a permanent virtual exhibition of research results, supplying to the platform new and improved versions of AI algorithms and other research outcomes to enhance the platform services. Include platform stakeholders into related dissemination and promotion campaigns	AILP owner, other developers as ‘exhibitors’
3	Organize common commercialization offer presentations, including meetings with AI-implementing companies. These activities may involve one or more of several consortium partners and are based on agreements regarding joint projects, arising as a direct consequence of such a meeting or another activity	All research partners
4	Cooperation of research institutions with industrial partners may be supported by different regional and national funds promoting industry-academia cooperation in AI. Such joint projects can only be supported if they have a commercial character. Diversified information campaigns should be addressed to public administration as well as to industry. Successful industry-academia cooperation will increase the chances of obtaining research funding by academic partners	All developer consortium members

during the roadmapping sessions during the last six months of the project performance period. Publicly available statistics, results of statistical forecasting and other IT/AI foresight exercises were also used.

Another characteristic feature of roadmapping-based planning is the need to take into account multiple conflicting criteria that describe the financial, technological and social goals of the planning object. These criteria are then applied to define the

set of non-dominated future plans. In the next step, multicriteria analysis and the corresponding decision support procedures were applied by roadmapping experts to point out several compromise solutions and select one for real-life implementation. This process is dynamic and the compromise plan, once chosen, can be updated to respond to varying external or platform-dependent circumstances. The collaboration roadmapping methodology complemented the AN-based model [10] when establishing a strategic plan for the above-presented AILP.

Finally, the provision of new content and services on the platform is modelled as an innovation development and market placement problem (ID-MP, cf. the New Product Development problem in [12]). The latter is a dynamic four-criteria problem with options-enhanced net present value (ENPV) [12], used as the principal investment selection criterion. ENPV aggregates subordinated momentary financial performance criteria. The remaining three quantitative indicators are:

- option-affected yield risk,
- social impact index (SII) defined as an aggregation of social impact measures proposed in [11],
- the Strategic Position Index (SPI) [12].

The multicriteria optimization problem that arises can be solved during an interactive group decision procedure with the above roadmapping methodology. The solution process can be assisted by an interactive real option detection algorithm, standardized real option valuation, collaborative SWOTC, Delphi-based investment efficiency assessment by experts and multicriteria analysis methods.

Social impact analysis is a supplementary activity that overlaps in part with the business-oriented roadmapping process. It refers to the micro- and meso-scale of platform uses. Social impact assessment and optimization are dealt with via other methods such as cellular-automata-based diffusion modelling [11]. Macro-scale social impact can be modelled by ANs and merged with the outcomes of the roadmapping process.

4 Summary and Conclusions

Marketing action planning as part of a platform's exploitation strategy supplements management and fosters service provision. The marketing of research results and of the platform as a whole may refer to research and innovation support policies at different levels, from regional to supranational. Building viable open-access AILP user communities may be regarded as a first step towards a distributed software development ecosystem, which comprises testing, debugging and user needs reporting. The research on AILPs, their marketing, social acceptance and governance is relevant to ensuring a positive impact and wide social support to AI strategies.

When designing the exploitation strategy and individual platform marketing activities, it turned out that existing approaches do not adequately suit the needs of planning the achievement of contractual goals that the developer consortium was

obliged to fulfil. The main goal of the platform has been the achievement of a prescribed number of users within the 5-year project durability period. We had to define a novel methodology of AILP exploitation strategy building with new paradigms and approaches, then involve all consortium partners and monitor the strategy implementation.

The origin of the roadmapping technique, which dates back to semiconductor production strategic planning in Motorola [14], then adopted in other large corporations, has left an imprint on the usual organization of this process, focusing on well-defined goals and involving the company's higher managerial staff. Collaboration roadmapping aims at coordinating diversified organizations from several countries, which follow different goals beyond common exploitation of the platform. Thus, when designing collaboration strategy, we faced new consensus finding problems that do not occur in single-company roadmapping processes. The newly designed collaborative RD building principles proved useful when fusing the outcomes of various forward-looking activities, including the expert Delphi and technological trend analysis, as well as taking into account consortium member preferences, market research and constraints imposed by external factors. Nevertheless, the efficient coordination of various goals in collaboration roadmapping is a challenge which requires further research and collection of real-life experiences in applying group decision support techniques and teleconferencing, rather than classical in-person sessions.

Beyond the technical aspects of marketing action planning, the exploitation strategy should often fulfil additional requirements of the grant-giving institutions, resulting in, for example:

- Conforming to regional- or national-level innovation strategies to ensure the achievement of sustainable development goals, based on smart specialization ranking. A real-life example: strategic planning for a regional innovation support centre [10], where recommendations to the R&D policymakers are derived from an anticipatory model.
- Selecting technological investment strategies for software companies, allowing corporate users to integrate the platform into their enterprise information systems.

The methodology presented in this paper can be regarded as a base for further research and development towards a general approach to building marketing strategies for information systems. Collaboration roadmapping, merged with other tools, such as causal and anticipatory networks, as well as with social impact simulation and optimization models, can become a universal market expansion planning methodology for web-based information systems and software-as-a-service (SaaS) platforms. The recommended decision models to selecting the ultimate exploitation strategy are those based on multicriteria analysis with ANs, reference sets and quantitative criteria describing the marketing goals achievement.

Acknowledgements This research was financed in part by the EC within the Horizon 2020 research project 'Training towards a society of data-savvy information professionals to enable open leadership innovation' (MOVING), contract No. 693092. The presentation of this research at the 3rd

ICMarkTech was supported by a grant from the Polish National Academic Exchange Agency (NAWA), contract No. PPI/APM/2018/1/00049/U/001.

References

1. Fenwick, D., Daim, T.U., Gerdtsri, N.: Value driven technology road mapping (VTRM) process integrating decision making and marketing tools: case of internet security technologies. *Technol. Forecast. Soc. Chang.* **76**(8), 1055–1077 (2009). <https://doi.org/10.1016/j.techfore.2009.04.005>
2. Fernandez-Flores, R. Hernandez-Morales, B.: Towards an open learning ecosystem based on open access repositories, curricula-based indices and learning management systems. In: Chova, L.G., Martinez, A.L., Torres, I.C. (eds.), 8th International Conference on Education and New Learning Technologies (EDULEARN), Barcelona, Spain, July 04–06, 2016—EDULEARN Proceedings, pp. 2826–2831 (2016)
3. Ho, C.K.Y., Ke, W., Liu, H.: Choice decision of e-learning system: implications from construal level theory. *Inf. Manage.* **52**, 160–169 (2015). <https://doi.org/10.1016/j.im.2014.07.003>
4. Kerr, C., Phaal, R.: Technology roadmapping: Industrial roots, forgotten history and unknown origins. *Technol. Forecast. Soc. Change* **155**, Art. No. 119967, 16 (2020), <https://doi.org/10.1016/j.techfore.2020.119967>
5. de Reuver, M., Sørensen, C., Basole, R.C.: The Digital platform: a research agenda. *J. Inf. Technol.* **33**(2), 124–135 (2018). <https://doi.org/10.1057/s41265-016-0033-3>
6. Ross, S., Hedstrom, M.: Preservation research and sustainable digital libraries. *Int. J. Digit. Libr.* **5**(4), 317–324 (2005)
7. Skulimowski, A.M.J.: Expert Delphi Survey as a Cloud-Based Decision Support Service, IEEE 10th SOCA 2017, 22–25 Nov. 2017, Kanazawa, Japan. IEEE, Piscataway, pp. 190–197 (2017), <http://ieeexplore.ieee.org/document/8241542>
8. Skulimowski, A.M.J.: Strategy building for a knowledge repository with a novel expert information fusion tool. In: 6th International Conference on Future-Oriented Technology Analysis (FTA): Future in the making, 4–5 June 2018, Brussels, pp. 1–15 (2018)
9. Skulimowski, A.M.J.: Forward-looking activities supporting technological planning of ai-based learning platforms. In: Herzog, M., et al. (eds.), *Advances in Web-Based Learning—ICWL 2019. Lecture Notes in Computer Science 11841*, Springer, Cham, pp. 274–284 (2019), https://doi.org/10.1007/978-3-030-35758-0_26
10. Skulimowski, A.M.J.: Multicriteria decision planning with anticipatory networks to ensuring the sustainability of a digital knowledge platform. In: Ben Amor, S. et al. (eds.), *Advanced Studies in Multi-Criteria Decision Making*. CRC Press, Boca Raton, pp. 168–197 (2020), <https://doi.org/10.1201/9781315181363-9>
11. Skulimowski, A.M.J.: User Community Development in Social Networks to Support AI-Enabled Knowledge Provision, 42nd International Conference on Information Systems (ICIS), Austin, TX, Proceedings, AIS eLibrary (2021), https://aisel.aisnet.org/icis2021/is_sustain/is_sustain/9
12. Skulimowski, A.M.J., Pukocz, P.: Enhancing creativity of strategic decision processes by technological roadmapping and foresight. In: Lee, V.C.S., Ong, K.-L. (eds.) *KICSS 2012: 7th International Conference on Knowledge, Information and Creativity Support Systems: Melbourne, Victoria, Australia, 8–10 Nov. 2012*. IEEE Computer Society, CPS, pp. 223–230 (2012), <https://ieeexplore.ieee.org/document/6405533>
13. Vagliano, I., Günther, F., Heinz, M., Apaolaza, A., Bienia, I., Breiffuss, G., Blume, T., Collyda, C., Fessel, A., Gottfried, S., Hasitschka, P., Kellermann, J., Köhler, T., Maas, A., Mezaris, V., Saleh, A., Skulimowski, A.M.J., Thalmann, S., Vigo, M., Wertner, A., Wiese, M., Scherp, A.: Open innovation in the big data era with the moving platform. *IEEE Multimedia* **25**(3), 8–21 (2018). <https://doi.org/10.1109/MMUL.2018.2873495>

14. Willyard, C.H., McClees, C.W.: Motorola's technology roadmap process. *Res. Manage.* **30**(5), 13–19 (1987). <https://doi.org/10.1080/00345334.1987.11757057>
15. Wu, D., Long, S., Zhou, X., Zhou, Y.: Research on marketing system construction of internet platform based on big data technology. In: 2021 International Wireless Communications and Mobile Computing (IWCMC), Harbin, China. IEEE, pp. 870–873 (2021), <https://doi.org/10.1109/IWCMC51323.2021.9498924>

Business Intelligence Databases and Marketing

Does Concept Recall in Brand Image Show High Loyalty? An Experimental Study on the Apple MacBook



Takumi Kato 

Abstract With technological advancements, product functionality and durability often exceed the standard that consumers require, and there is little difference between brands. This implies that competitiveness between brands is shifting to emotional value, such as design and usability, from functional value. The importance of emotional value is now recognized, and companies increasingly focus on the value creation. Despite improved design and usability, products may seem similar from the consumer's perspective, and it becomes difficult to enhance their perceptions of product value. This may be attributed to the ambiguous concept. Design and usability are effective means of embodying value, but these are not the essence of value. Consumers who attracted to superficial features like design will easily switch brands when another product with a more fashionable design appears. If a consumer remains loyal to a brand, it is because they understand, sympathize with, and value the brand's concept. However, few studies have focused on brand concept as a factor of loyalty. Accordingly, the present study hypothesized that consumers who recall a brand concept are more likely to exhibit loyalty than consumers who recall specific features, such as design and usability. It was evaluated the four brands loyalty factors—concept, design, usability, and technology—for Apple's MacBook, which has a strong brand in Japan. The results showed that the concept effect was the largest, proving that concept management is crucial for building and growing a strong brand. Hence, the concept recall index should be emphasized in brand management.

Keywords Brand concept · Emotional value · Design · Usability

1 Introduction

The market is full of products that have not developed into a strong brand and are buried among similar products. One of the root causes, concept of a product, is considered ambiguous. Concept is defined as values, consisting of target (who),

T. Kato (✉)

Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan
e-mail: takumikato@mail.saitama-u.ac.jp

value and positioning (what), and execution method (how) [1, 2]. For a product to become a strong brand, it is necessary to clearly define the concept before the product even enters the market, maintain the concept, and continue to embody it as perceived by the consumer [3].

In recent times, the source of competitiveness for products has shifted to emotional value, such as design and usability, from functional value [4]. Emotional value that meets psychological needs is the most important predictor of purchase behavior of the product/service [5]. Although corporate engineers were not previously interested in products' subjective esthetics [6], the importance of emotional value is now recognized, and companies increasingly focus on design and usability. However, despite improved product design, products may seem similar from the consumer's perspective, and it becomes difficult to enhance their perceptions of product value. Thus, whether a product can be established as a strong brand in the market depends on a well-defined attractive concept. Consumers may be attracted to superficial functionality and design, but they may easily switch brands when a different company's product offers higher functionality and a more fashionable design. A consumer's brand loyalty stems from how well they understand, sympathize with, and value the brand's concept [7].

However, there are few examples of this concept as a factor of loyalty. Although this concept has been addressed many times in marketing and business administration-related discussions, its effect has not been quantitatively demonstrated. Therefore, this study comprehensively verified the contribution of the four factors of concept, design, usability, and technology toward customer loyalty for a personal computer brand, Apple MacBook Pro (hereinafter, MacBook), in the Japanese market. The study also provides suggestions for both business and research to reaffirm the importance of the concept as a brand loyalty factor.

2 Literature Review and Derivation of Hypothesis

Product concept is the criterion for decision-making in all corporate activities, such as planning, development, production, and sales [8, 9]. Product development involves many daily decisions by various departments, and consistent standards must be applied throughout the process. Additionally, the concept significantly expands a brand or helps to form brand alliances. The distinguishing factor between success and failure of a brand extension is consistency in categories and concepts [10–12]. Consumers who exhibit high loyalty also have high demands for consistency [13], and the consistency of the concept is more important than its category.

Despite the big role of concept in a product's journey, few studies have evaluated it as a factor of loyalty. Some concepts are divided into general categories (e.g., functional and symbolic) [14–17]. However, it is difficult to determine whether these methods highlight consumers' understanding of product concepts because the general categories are too broad. There are three possible reasons why the concept has not been addressed. First, researchers and corporate marketers believe it is difficult for

consumers to evaluate concepts that cannot be directly experienced. Second, many product concepts are often ambiguous, and marketers can lose sight of their purpose because they cannot clearly identify how their products/services are meaningful to customers [18]. When the concept is ambiguous, products rely on trendy, superficial designs, and advanced functions, and it is difficult to evaluate the brand concept. Third, even if there is a clear concept, there may not be a consistent embodiment of the product, and it comes across as ambiguous in the marketplace. Much effort is required to consistently embody a concept without compromises when each corporate department makes different claims; this is the reason why a thoroughly implemented product that can grow into a strong brand is a rare entity.

Consumers form brand images through their experiences with the product brand. Whether consumers recall the product concept when asked about the brand image is expected to have a significant impact on their loyalty. Therefore, when asked about the same, consumers should be able to recall the keywords in their concept. This is because a consumer's brand loyalty stems from how well they understand, sympathize with, and value the brand's concept [7]. If the focus is on the product with a clear concept that builds a strong brand in the market, evaluation becomes possible. Accordingly, the following hypothesis was formulated.

H1: Recalling concepts as the brand image positively impacts brand loyalty.

Next, from the perspective of products, the influence of emotional value has been pointed out recently [19]. In particular, concept [20], design [21–24], and usability [25–28] have been demonstrated and claimed to be effective in many studies. For example, in the past, product sound design had been aimed at noise reduction; however, recently, its purpose has changed to convey the product concept and provide emotional experience [29]. Accordingly, the following hypothesis was formulated.

H2: Recalling designs as the brand image positively impacts brand loyalty.

H3: Recalling usability as the brand image positively impacts brand loyalty.

However, functional value is never unnecessary. Regardless of how good the design is, the product will not be accepted by consumers if the quality or safety performance is defective. Many emotional values are generated from corresponding functional values [30]. For example, comfort levels in the interior of an automobile are realized by the quietness of the engine and the suppression of vehicle body vibration. In addition, among the functional values, technology provides consumers with direct value [31–33]. Moreover, in order to effectively convey the technology to consumers, the technology should be given a name and promoted as a brand [34, 35]. This is because a brand has a much greater effect than the characteristics of its object [36]. In fact, many have shown that brand logos have a greater impact on consumer behavior than the objective characteristics of the subject [37–39]. Accordingly, the following hypothesis was formulated.

H4: Recalling technologies as the brand image positively impacts brand loyalty.

H5: Recalling brands as the brand image positively impacts brand loyalty.

And finally, as already mentioned, it is presumed that consumers with higher loyalty place more emphasis on the concept that is the essence of the product, and the following hypothesis was derived.

H6: Consumers who recall a brand concept are more likely to exhibit loyalty than consumers who recall specific features, such as design and usability.

3 Method

3.1 Survey

An online survey was conducted in Japan from November 5 to 10, 2020. The respondents' eligibility criteria were: (a) age ranging from 20 to 59 years, (b) MacBook owner, and (c) use MacBook at least once a month. The survey comprised questions regarding: (1) gender, (2) age, (3) use frequency, (4) preference, (5) recommendation intention, (6) repurchase intention, and (7) brand image. From among the responses to (1) and (3), those from people who did not meet the eligibility criteria were automatically excluded, and 400 responses were finally collected. Additionally, responses to gender and generation were evenly distributed. Loyalty indices (4)–(6) were rated on a 7-point Likert scale (1 = not at all preferable and 7 = extremely preferable). The mean values for each are as follows: preference, recommendation, repurchase = 5.928, 5.523, 5.713, respectively. Compared to the other two indicators of self-emotion, recommendation intention, which requires involvement with others, is the answer with the least value. This means that it has the highest psychological barriers and is therefore frequently used as a loyalty indicator. The question pertaining to brand image was asked because, as mentioned in Sect. 2, consumers form brand image through brand experience. Item (7) was a pure recall question, wherein respondents were not presented with options. This is because aided recall, which presents options, introduces bias, and the options may then be overestimated. By using pure recall, the respondent could provide an answer about the concept only if they really understood it.

3.2 Verification

In this study, natural language processing was used to extract words related to concept, design, usability, technology, and brand from the text of the responses to brand image. Since many expressions such as “brand power” and “high brand” were seen, brand factors were also set. To exclude subjectivity, five words belonging to each factor were set. As shown in Table 1, words belonging to each factor were defined. The concept of MacBook as extracted from the official Web site is: “the ultimate work tool for professionals who innovate in the world” [40], so target words were

Table 1 Words in each factor and number of detections of each factor

Word	Concept	Design	Usability	Technology	Brand
1	Ultimate	Design	Usability	Technology	Brand power
2	Work tool	Looks	Operability	High performance	High brand
3	Professional	Appearance	User-friendly	Functionality	Philosophy
4	Innovative	Stylish	Stress-free	Powerful	Name value
5	Creative	Fashionable	Intuitive	Speed	Reputation
Number of detected words in each factor	44	162	67	114	43
Number of mention flags for each factor	35	127	59	94	35

set based on this concept. In Japanese, there are four ways to express an idea in writing: kanji, hiragana, and katakana for Japanese words, and katakana for English words. Different notations for the same words were extracted similarly. Therefore, the words (nouns and adjectives) of interest, including the different notations for the same words, were extracted from the respondents’ sentences using morphological analysis, and the detected words were marked with mention flags. The Japanese open-source software, MeCab, was used. Table 1 shows that 430 words were extracted from the 400 responses, and 350 mention flags were identified for all the factors combined. This implies that respondent mentions of the same factor were not counted twice, and each factor was represented by its respective mention flags.

As for the image of MacBook, design was found to be the most recalled factor, followed by technology and usability. Although the recall frequency is centered on the features that come into direct contact with consumers, the interest of this study was its effect on loyalty. As shown in Table 2, the variables for the evaluation were loyalty indices (preference, recommendation, and repurchase) and factors (concept, design, usability, technology, and brand). Using these variables, the hypotheses were

Table 2 Variable list

No.	Classification	Variable	Type	Mean	SE
1	Loyalty index	Preference	7-point scale	5.928	0.057
2		Recommendation	7-point scale	5.523	0.060
3		Repurchase	7-point scale	5.713	0.061
4	Recalling dummy	Concept	0/1	0.088	0.014
5		Design	0/1	0.318	0.023
6		UX	0/1	0.148	0.018
7		Technology	0/1	0.235	0.021
8		Brand	0/1	0.088	0.014

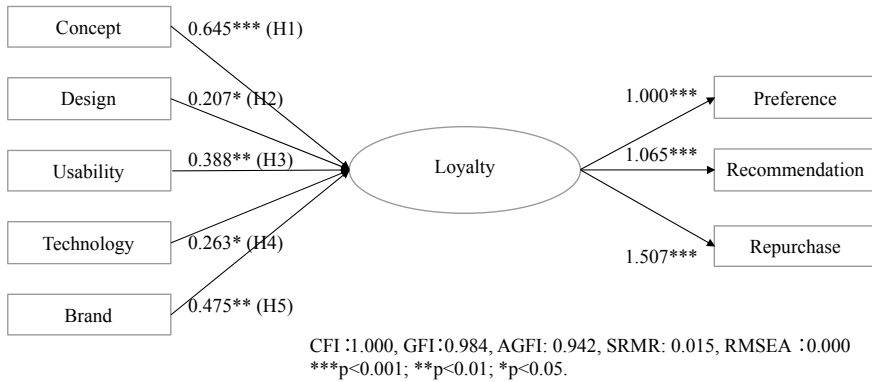


Fig. 1 Result of the structural equation modeling

tested using structural equation modeling. The analysis environment was R, and the lavaan package was used for structural equation modeling.

4 Results

Figure 1 shows the results of the structural equation modeling. The indicators of the model showed high suitability: CFI = 1.000, GFI = 0.984, AGFI = 0.942, SRMR = 0.015, and RMSEA = 0.000. All five factors had a significant effect on loyalty; hence, the hypotheses H1, H2, H3, H4, and H5 were supported. Furthermore, the factor that has the greatest influence on loyalty is the concept, followed by brand, usability, technology, and design. Comparing Table 1 and Fig. 1, the superiority and inferiority of frequency and contribution are different. Therefore, decision-making should not be distorted by being overwhelmed by numerous voices and trends. Thus, the hypothesis H6 is supported.

5 Practical Implications and Limitations

Regarding practical implications of this study, companies should first reaffirm the importance of brand concepts. There are still many products/services equipped with superior functions and designs, while their brand concept remains ambiguous. Accordingly, the concept recall index should be emphasized in brand management. The index is used to evaluate whether or not the concept is recalled when the consumers are asked about brand image. Like many previous studies, this study also found that design, usability, and technology certainly contribute to loyalty. However, the effect of concept was greater than all others. Hopefully, the index of concept

recall will gain prominence not only in the industrial world, but also in the academic world that studies consumer behavior and brand management.

This study has three limitations. First, since it covered only Japan, the generalizability of its results to other countries remains limited. Second, the results may vary depending on the evaluation method of concept recall. In this study, pure recall was adopted to eliminate subjectivity, and five words were used for detection from sentences. This index should be continuously examined in the future. Third, since pure recall was applied, features that consumers usually have difficulty recalling are unlikely to appear in the evaluation (e.g., corporate social responsibility, consideration for the environment). Therefore, only typical features, such as design and usability, were compared in this study. In other words, the survey method changes depending on whether the evaluation is based on factors that are easy for consumers to notice without bias or on factors that are difficult for consumers to notice despite bias. In brand management practice and research, a discussion on the method of evaluating concepts is still lacking, and further research is required.

References

1. Lafley, A.G., Martin, R.L.: *Playing to Win: How Strategy Really Works*. Harvard Business Press, Boston (2013)
2. Stengel, J.R., Dixon, A.L., Allen, C.T.: Listening begins at home. *Harv. Bus. Rev.* **81**(11), 106–117 (2003)
3. Gardner, B.B., Levy, S.J.: The product and the brand. *Harv. Bus. Rev.* **33**(2), 33–39 (1955)
4. Noble, C.H., Kumar, M.: Using product design strategically to create deeper consumer connections. *Bus. Horiz.* **51**(5), 441–450 (2008)
5. Hur, W.M., Kim, Y., Park, K.: Assessing the effects of perceived value and satisfaction on customer loyalty: a ‘Green’ perspective. *Corp. Soc. Responsib. Environ. Manag.* **20**(3), 146–156 (2013)
6. Adams, J.: *Good Products, Bad Products: Essential Elements to Achieving Superior Quality*. McGraw-Hill Education, New York (2011)
7. Aaker, D.A., Joachimsthaler, E.: *Brand Leadership*. The Free Press, New York (2000)
8. Tilley, C.: Built-in branding: How to engineer a leadership brand. *J. Mark. Manag.* **15**(1–3), 181–191 (1999)
9. Simões, C., Dibb, S.: Rethinking the brand concept: new brand orientation. *Corporate Commun.: Int. J.* **6**(4), 217–224 (2001)
10. Jin, L., Zou, D.: Extend to online or offline? the effects of web-brand extension mode, similarity, and brand concept on consumer evaluation. *J. Mark. Manag.* **29**(7–8), 755–771 (2013)
11. Lanseng, E.J., Olsen, L.E.: Brand alliances: the role of brand concept consistency. *Eur. J. Mark.* **46**(9), 1108–1126 (2012)
12. Park, C.W., Milberg, S., Lawson, R.: Evaluation of brand extensions: the role of product feature similarity and brand concept consistency. *Journal of Consumer Research* **18**(2), 185–193 (1991)
13. Samuelsen, B.M., Olsen, L.E., Keller, K.L.: The multiple roles of fit between brand alliance partners in alliance attitude formation. *Mark. Lett.* **26**(4), 619–629 (2015)
14. Bhat, S., Reddy, S.K.: Symbolic and functional positioning of brands. *J. Consum. Mark.* **15**(1), 32–43 (1998)
15. Hagtvedt, H., Patrick, V.M.: The broad embrace of luxury: hedonic potential as a driver of brand extendibility. *J. Consum. Psychol.* **19**(4), 608–618 (2009)

16. Jeon, J.E.: The impact of brand concept on brand equity. *Asia Pacific J. Innov. Entrepreneurship* **11**(2), 233–245 (2017)
17. Park, C.W., Jaworski, B.J., MacInnis, D.J.: Strategic brand concept-image management. *J. Mark.* **50**(4), 135–145 (1986)
18. Blount, S., Leinwand, P.: Why are we here? *Harv. Bus. Rev.* **97**(6), 132–139 (2019)
19. Kato, T.: Synergistic effect of matching corporate and product brand images on purchase intentions: comparing the importance of functional and emotional value. *J. Brand Manag.*, 1–14 (2021)
20. Kato, T.: Contribution of concept recall to brand loyalty: an empirical analysis of design and performance. *J. Consumer Behaviour*, 1–10 (2021)
21. Homburg, C., Schwemmler, M., Kuehnl, C.: New product design: concept, measurement, and consequences. *J. Mark.* **79**(3), 41–56 (2015)
22. Hsu, C.L., Chen, Y.C., Yang, T.N., Lin, W.K., Liu, Y.H.: Does product design matter? exploring its influences in consumers' psychological responses and brand loyalty. *Inf. Technol. People* **31**(3), 886–907 (2018)
23. Kim, C., Christiaans, H.H.: The role of design properties and demographic factors in soft usability problems. *Des. Stud.* **45**, 268–290 (2016)
24. Orth, U.R., Malkewitz, K.: Holistic package design and consumer brand impressions. *J. Mark.* **72**(3), 64–81 (2008)
25. Babbar, S., Behara, R., White, E.: Mapping product usability. *Int. J. Oper. Prod. Manag.* **22**(10), 1071–1089 (2002)
26. Lee, D., Moon, J., Kim, Y.J., Mun, Y.Y.: Antecedents and consequences of mobile phone usability: linking simplicity and interactivity to satisfaction, trust, and brand loyalty. *Inf. Manag.* **52**(3), 295–304 (2015)
27. Mugge, R., Schoormans, J. P.: Product design and apparent usability. The influence of novelty in product appearance. *Appl. Ergon.*, **43**(6), 1081–1088 (2012)
28. Seva, R.R., Gosiaco, K.G.T., Santos, M.C.E.D., Pangilinan, D.M.L.: Product design enhancement using apparent usability and affective quality. *Appl. Ergon.* **42**(3), 511–517 (2011)
29. Dal Palù, D., De Giorgi, C., Lerma, B., Buiatti, E.: *Frontiers of Sound in Design: A Guide for the Development of Product Identity Through Sounds*. Springer, Cham (2018)
30. Ghodeswar, B.M.: Building brand identity in competitive markets: a conceptual model. *J. Product Brand Manage.* **17**(1), 4–12 (2008)
31. Kimiloğlu, H., Nasır, V.A., Nasır, S.: Discovering behavioral segments in the mobile phone market. *J. Consum. Mark.* **27**(5), 401–413 (2010)
32. Van Wezemael, L., Ueland, Ø., Rødbotten, R., De Smet, S., Scholderer, J., Verbeke, W.: The effect of technology information on consumer expectations and liking of beef. *Meat Sci.* **90**(2), 444–450 (2012)
33. Vyas, V., Raitani, S.: Cross-buying as a challenge in emerging markets: a study of the effect of technology adoption on cross-buying intentions. *Int. J. Finan. Innov. Bank.* **1**(1–2), 80–98 (2016)
34. Liu, P., Fei, Q., Liu, J., Wang, J.: Naming is framing: the framing effect of technology name on public attitude toward automated vehicles. *Public Understanding of Science*, 0963662520987806 (2021)
35. Kato, T.: Comparison of the influence of self-driving technology brand name on purchase intention between Japan and the US. *Procedia Manuf.* **39**, 1369–1376 (2019)
36. Aaker, D.: *Aaker on Branding: 20 Principles that Drive Success*. Morgan James Publishing, New York (2014)
37. McClure, S.M., Li, J., Tomlin, D., Cypert, K.S., Montague, L.M., Montague, P.R.: Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**(2), 379–387 (2004)
38. Brasel, S.A., Gips, J.: Red Bull “gives you wings” for better or worse: a double-edged impact of brand exposure on consumer performance. *J. Consum. Psychol.* **21**(1), 57–64 (2011)

39. Fitzsimons, G.M., Chartrand, T.L., Fitzsimons, G.J.: Automatic effects of brand exposure on motivated behaviour: how apple makes you “think different.” *Journal of Consumer Research* **35**(1), 21–35 (2008)
40. Apple.: MacBook Pro. Apple, <https://www.apple.com/jp/macbook-pro-16/> (n.d.)

What Makes a Movie Get Success? A Visual Analytics Approach



Bruno Vaz, Maria de Fátima Barros, Maria João Lavoura,
and Álvaro Figueira 

Abstract It is common for people to choose their next movie or show through other viewers' experience statements, like the Internet Movie Database (IMDb) presents. In this paper, we will be inspecting the IMDb public datasets, processing them, and using a visual analytics approach to understand how a movie can be successful among its fans. The main exploration focus is regions where titles are translated to, how the success of a title relates to its cast, crew, and awards nominations/wins. We took a methodology based on hypothesis formulation based on the EDA exploration and their testing based on a visual analytics confirmation.

Keywords Visual analytics · IMDb · Movie success · Data analysis

1 Introduction

The Internet Movie Database (IMDb)¹ is a well-known source for obtaining information about titles, TV shows, actresses, actors, directors, screen players, and much more. People are accustomed to selecting their next film or program according to its ratings and reviews. This database comprehends an immense opportunity to relate information. For instance, we have an insight of the crew members with details on some information relevant to their lives [1] but also to other movies or directors. It also specifies different labels for each instance according to world region and even describes the duration and genres of each title.

¹ <https://www.imdb.com/pressroom/>

B. Vaz · M. J. Lavoura · Á. Figueira (✉)
Faculty of Sciences, University of Porto, Rua Campo Alegre s/n, 4169-007 Porto, Portugal
e-mail: arfiguei@fc.up.pt

M. de Fátima Barros
Faculty of Engineering, University of Porto, Rua Roberto Frias, 4200-465 Porto, Portugal

Á. Figueira
INESC TEC, Rua Roberto Frias, 4299-465 Porto, Portugal

Table 1 Brief description of IMDb datasets and variables used

Dataset	Description
<i>Title Basics</i> 5,752,893 rows	This dataset contains the main information for each title. That is, the name of movie or show. It includes features such as the identity of the title (<i>tconst</i>), the most used title (<i>primaryTitle</i>), the title in the original language (<i>originalTitle</i>), the genres (<i>genres</i>) and if it's for adults (<i>isAdult</i>). It also includes the year it was released (<i>startYear</i>), and for TV series, the end year (<i>endYear</i>)
<i>Title Akas</i> 19,922,791 rows	In this dataset, according to the <i>titleId</i> (of type <i>tconst</i>), we can find enumerated information (<i>ordering</i>) about the title (<i>title</i>) by region (<i>region</i>) and language (<i>language</i>), taking in consideration if it is the original title or not (<i>isOriginalTitle</i>). A set of numerated attributes (<i>types</i>) and not numerated (<i>attributes</i>) is also provided
<i>Title Ratings</i> 1,086,028 rows	For each title (<i>tconst</i>), we can find in this dataset the corresponding rating (<i>averageRating</i>), and number of votes (<i>numVotes</i>)
<i>Title Crew</i> 7,281,233 rows	This table refers to the directors (<i>directors</i>) and writers (<i>writers</i>) for each title (<i>tconst</i>)
<i>Title Principals</i> 40,449,024 rows	In each title (<i>tconst</i>), principal cast and crew are listed (<i>ordering</i>) and detailed with an Id (<i>nconst</i>), describing their job category in a given title (<i>category</i>), and specific job, if applicable (<i>job</i>). For actors, we may also find the name of the character played (<i>characters</i>)
<i>Name Basics</i> 6,283,772 rows	IMDb also provides information about titles' celebrities. So, in this dataset, for each person (<i>nconst</i>), their name (<i>primaryName</i>), birth year (<i>birthYear</i>), primary professions (<i>primaryProfession</i>), and titles the person is known for (<i>knowForTitles</i>) are described. The death year (<i>deathYear</i>) is also provided, if applicable

In this paper, we analyze the IMDb dataset,² which is publicly available and updated daily (data retrieved in October 2020). Our motivation is to find reasons behind the success of movies, that can be identified and quantified, which ultimately can then be used to obtain more profit from the movies enlarging the fan base, and offer to the public what rationally (or not) is wanted. For such, first, a description of the dataset is presented. Then, we review the processing techniques used for filtering and cleaning the data, making it useful for the visual analysis [2]. Then we formulate hypothesis concerning the movies' success. At last, we explore and prove (or refute) the proposed hypothesis. All figures in this paper were created using R and plotly [3].

2 The Datasets

In this research, we used 6 of the 7 available IMDb data tables (in CSV). We excluded the one with information about TV Series or episodes, which is not our focus. In Table 1, a brief description of the information for each table is provided.

² <https://www.imdb.com/interfaces/>

2.1 Filtering

Table *Title Basics* was used to filter the data through the type of title, including the categories *short*, *movie*, and *tvMovie*, and excluding the remaining (related to tv series and video games) obtaining now 1,443,759 rows. After this, the filtering process for *Title Akas*, *Title Principals*, *Title Crew*, and *Title Ratings* was based on the title Id's (*tconst*) present in the *Filtered Title Basics*. Then, for filtering *Name Basics*, we used the instances of people's Id's in the now filtered *Title Principals*.

2.2 Cleaning

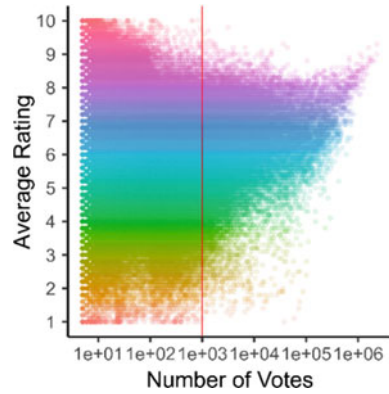
Title Basics: Entries without a start year were removed (6.1%). The column *endYear* only referred to series, so it was deleted. Entries with the duration (*runtimeMinutes*) of more than 4 h were also removed since they are considered to be outliers and only represent around 0.067% of the instances. Finally, the instances without any value for the attribute *genres* (around 5.7%) were removed. Since the column *genres* have up to 3 values, it was divided into 3 columns each one containing only one genre and a 0 whenever there is no data. The cleaned table was left with 1,277,962 rows.

Title Akas: After filtering has 2,211,001 rows. The *ordering* feature does not add any additional information, so it was erased. For the *attributes* and *types* features, most of the entries do not have this information (95 and 50%), so they were discarded. The localized title (*title*) does not include additional information, and it was problematic since it comes in different types of alphabets. Hence, it was removed. For *region* feature, 13.6% of the entries do not have a value, so they were also removed. *Language* feature also has much missing data (85.1%), so this column was removed. The *isOriginalTitle* feature has, now, 0.0016% of missing values, being also removed. At last, after cleaning, *Title Akas* was left with 1,911,339 entries (a decrease of 13.6%).

Title Ratings: After filtering, the *Title Ratings* table was left with 423,183 rows, these corresponding to *movie*, *tvMovie*, and *short*. We can state there are no *averageRatings* outside the interval $[0,10]$, nor are negative values for the attribute *numVotes*. Since there are also no missing values, no preprocessing was made for this table. Besides this and analyzing the distribution of average ratings versus the number of votes (Fig. 1), we decided to only consider entries with *numVotes* bigger than 1000.

Title Crew: The filtering process left the *Title Crew* table with 1,443,759 instances. We noticed there were around 7.8% observations with neither a director nor a writer. Since this is not useful, they were removed. Then, two new columns were created, one containing the number of directors per instance and the other the number of writers, also per instance. The table was left with 1,330,669 rows. Additionally, we created a new table from *Title Crew* transforming each column of writers and directors into rows, so each instance contained a *tconst*, a *nconst*, and a boolean for determining if the *nconst* referred to either a director or a writer.

Fig. 1 Scatter plot of number of votes vs. average ratings (threshold of 1000 vote in red)



Title Principals: In *Title Principals*, we have 4,621,598 instances (post-filtering). The *ordering* column did not give any relevant information, so we discard it. Also, 85.5% and 62.7% of the population did not have a *job* or *characters*, respectively, so these features were also discarded. After this, we were left with the variables *tconst* for identifying the movie, *nconst* for the person in question, and their *category*. Since no rows were removed, we remained with the starting number of instances.

Name Basics: After filtering, the table had 1,634,635 instances. We discovered that 85% of these did not have a birth year entrance, so we decided to discard this variable. As for the death year, we transformed the data into a new column *isDead*, a Boolean. Considering that each person may have up to 3 professions, the *PrimaryProfession* column was replaced by 3 new columns, each with only one value, where empty entries are imputed with 0's. We removed empty instances (6.3%) of the *knownForTitles* feature, decreasing the rows to 1,532,134. The *knownForTitles* attribute has a frequency varying between 1 and 6 values per row. However, since only 23 rows have 5 different values of *knownForTitles* and only 2 rows have 6 different values, we discarded these and replaced it by 4 new ones with the first 4 *knownForTitles* values.

2.3 Merging

After filtering and cleaning, the *Title Basics*, *Title Crew*, and *Title Ratings* datasets were left joined by the *tconst* attribute. We left joined *Title Basics* and *Title Crew*. Then, we left joined the resulting dataset and *Title Ratings*. The resulting merged dataset has some instances in which the values of the attributes of the *Title Crew* and *Title Ratings* datasets are missing. In such cases, a global constant (−1) was imputed.

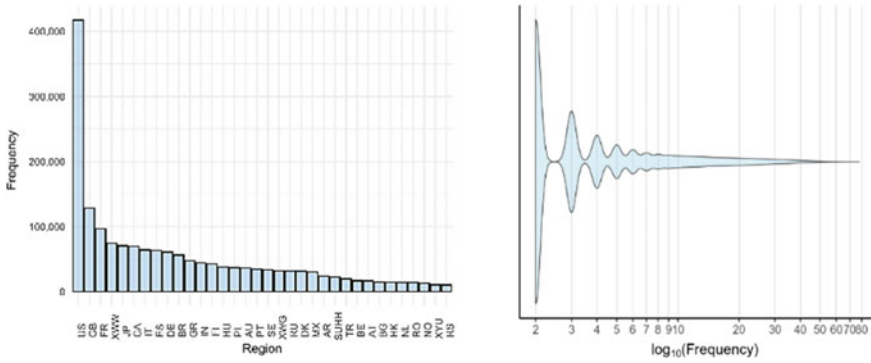


Fig. 2 Number of titles per region (left). Distribution of unique titles with two or more translations (right)

3 Exploratory Data Analysis and Hypothesis Formulation

Table *Title Akas* induced us to study the distribution of titles around the globe. In Fig. 2 (left), we plot a frequency distribution of translated titles to different regions. It is noticeable that the USA has a much more of translated titles than any other region. Moreover, for each title, only 25% are translated to other languages. We depicted the distribution of the regions per title of these 25% in a violin plot Fig. 2 (right).

From the plot, we understand that most of the titles are only translated to two or three regions, but some of them are translated to dozens of other languages, relating to other regions in the world. Hence, we created the following hypothesis: “Is the title’s success (rating and number of votes) higher when it is translated to other languages?”.

In addition, the average ratings of titles per person, which results from the combination of *averageRating* and *knownForTitles* features, in *Title Ratings* and *Name Basics*, respectively, are also considered average, as we can inspect when analyzing Fig. 3 (right). Hence, we are led to formulate the following question: “What makes a movie successful for the voters?”.

We wanted to go beyond the IMDb dataset to inspect the relationship between titles and their success by analyzing the movie awards. We used three different datasets publicly available: the *Emmy Awards* dataset³ contains data from 1949 up to 2020. The *Golden Globe Awards*, 1944–2020 dataset,⁴ and finally, the *Oscar Awards*, 1927–2020⁵. We intersected these datasets with *Title Basics* dataset through the name of titles and their premier year. The analysis of this dataset made us question if the titles are more recognized for the motion picture itself, or by the people that make the movie. More than ¾ of the nominees in the *Golden Globes* dataset are people, and not

³ <https://www.kaggle.com/unanimad/emmy-awards>.

⁴ <https://www.kaggle.com/unanimad/golden-globe-awards>.

⁵ <https://www.kaggle.com/unanimad/the-oscar-award>.

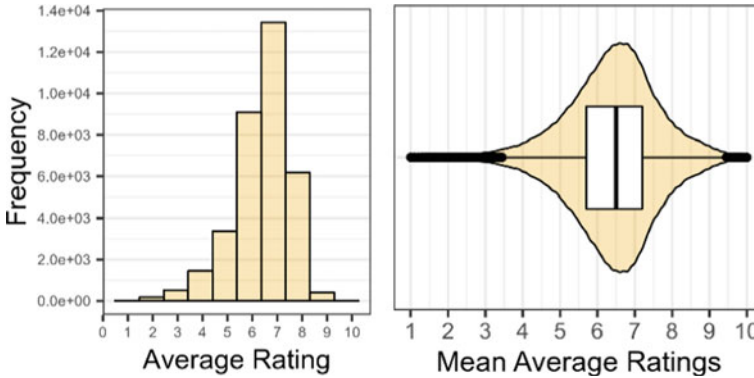


Fig. 3 Average ratings per title/person

titles since most categories are made for the crew and cast. Since most movie awards categories rely on the people making the movie, hypothesize: “Does the cast/crew of a movie make it successful?”, when it comes to ratings, votes, and even awards.

4 Hypotheses Assessment

4.1 Titles Per Region

To have a general overview regarding the distribution of the titles concerning the countries they were translated we plot a world map (using, among other, techniques described in [3] with circles in the regions present in table *Title Akas*, and whose area is given by the number of titles translated to that country. Moreover, the circles have different colors representing the most frequently translated genres.

First, we remove some regions from the *Title Akas* dataset, for they do not represent countries (e.g., XWW, which means worldwide; XEU, represents the European continent). These cases represent about 8% of the rows. We also used a dataset *Countries*⁶ with coordinates (latitude, longitude) representing each country. We then merged *Title Akas* with the *Merged Basics Ratings Crew* by *tconst* and transformed the resulting dataset, so each region had one and only one row, its coordinates and the number of titles for each of the genres. The resulting plot is represented in Fig. 4 (left). By analyzing the graphic, we see that the USA is the country to which most titles are translated to, and the predominant genre is drama. But drama is not only the most frequent genre in the USA. As one can notice from the number of yellow dots, this genre is the most translated worldwide. We can also observe that the more developed countries tend to have more titles translated to. This type of graphic led

⁶ <https://www.kaggle.com/paultimothymooney/latitude-and-longitude-for-every-country-and-state>.

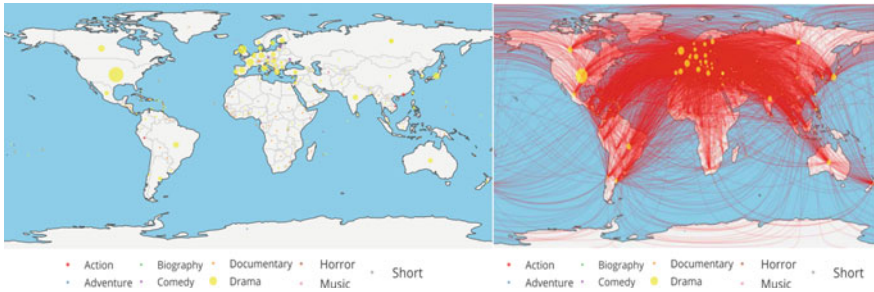


Fig. 4 Most frequent genres (left). Connections between countries and most frequent genres

us to the idea of representing the links between regions. That is, given a movie, we assess to which regions it was translated to and we connected them by an edge, 2 by 2.

From the *Title Akas*, we removed countries only translated to one region, since they wouldn't be useful for our purposes and we transformed the resulting dataset, so each movie had one and only one row, the regions to which it was translated to and the number of such regions. Then we have a table of connections between regions and the frequency of these connections. That is, each row represents a unique connection: as attributes we had two regions and the number of links between them (the number of titles that were translated to both regions) as depicted in Fig. 4 (right). As we can see, the titles that are translated to more than one region are typically translated to some European country.

We analyze now how the number of regions and the mean average ratings correlate. Using *Title Ratings*, we can, for each region, make the mean of the *averageRatings* of all the titles that were translated to that region, and create the graphic in Fig. 5 where the area of the circles represents the number of titles that were translated to a certain number of regions. The big blue circle represents the number of titles translated to

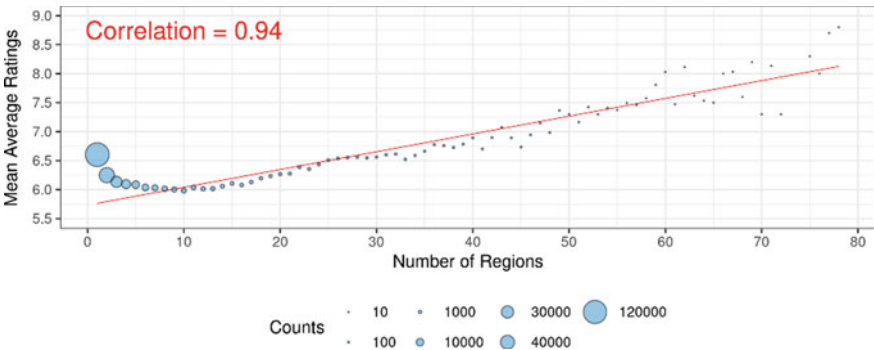


Fig. 5 Correlation between number of regions and mean average ratings

only one region, which is the majority. As the number of regions increases fewer titles are translated to that many regions, but the mean of their ratings increases.

We can see the correlation is very close to 1, meaning that titles with higher ratings are translated to more regions. This gives a very clear answer to our question. Indeed, the success of a movie is higher when translated to other languages.

4.2 Title's Success

Cast and Crew Influence: To explore the hypothesis “what makes a movie successful?” we focused on answering more specific questions related to the hypothesis like “Does the cast of a movie influences its success?” and “Is there a specific profession that influences the success of a movie?”.

To answer these questions, we started by defining “success”. We took a slightly different approach from [4] and looked into the distribution of the number of votes (Fig. 1) and the histogram of the average ratings (Fig. 3, left). Since the rate range is 1 to 10, we define a successful movie as having an average rating ≥ 7 and an unsuccessful is one that has an average rating ≤ 3 . We also filtered the *Merge Basics Ratings Crew* with instances that had *numVotes* > 1000 .

Usually, the most commonly known cast professions for a movie are actors, actresses, directors, and writers. To know how a cast influences a movie's success, we filtered the *Name Basics* keeping at least one of the three columns of primary professions equal to one of the above mentioned. We realized that there are some huge values regarding the number of people that are known for a specific movie. People being known for a single movie goes from 1 to 3087. Even though the whole crew of a movie can be as big as 3087, it is difficult to accept that the whole crew gets to be known by it. Moreover, for visualization purposes, we filtered it by the top 10 frequencies (Fig. 6). We also looked only at the cast that is still alive.

Fig. 7 is represented a graph where each node is a movie, and each link is a casting person in common. The red nodes represent unsuccessful titles, and the green nodes represent successful titles. The links are colored based on the primary profession of each person, with actor/actress being red, writers being blue, directors being black,

Fig. 6 Top 10 frequency of known actors (goes up to 3087 people)

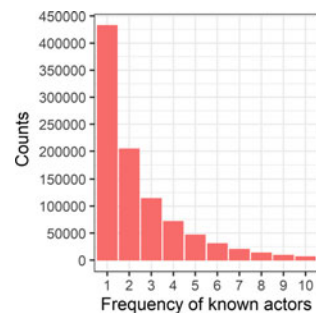
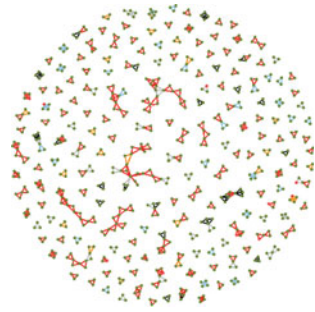


Fig. 7 Relationship between cast and crew of the titles they participate in



producers being orange and other professions being gray. To analyze titles that have more connections between them, a filter was applied where a person has to be known for 3 or 4 titles. This graph is only represented 11 unsuccessful titles against 753 successful ones. There many of smaller groups of 3, 4, 5, 6, 7, or 8 titles bonded together, some median chains bonded by 9, 10, 12, 13, or 15 titles, and 2 bigger chains of 21 and 27 titles connected. The unsuccessful titles are present in small size chains (groups of totals of 3, 4, 5, 6, or 8 titles connected). The successful titles are present in all sizes of chains. Therefore, we can conclude that a successful movie tends to have more common people with other successful titles, than with unsuccessful ones. Regarding the links in the graph, there is a clear majority of actors and actresses, followed by other professions and directors, writers and lastly producers.

The 5–8 node chains are connected by two different link colors, so 2 different professions. The medium and big chains are the ones that gather up to three different professions within their connections. The bigger chains (21 and 27 titles) and medium chains (from 9 to 15 titles) are dominated by actors/actresses. Therefore, we can conclude that the most relevant profession in successful titles is indeed the actor/actress.

We explored the relationships among the crew (directors and writers) within a title and their relationships in successful titles, by merging the contents from *Title Crew* and *Title Ratings* and making a tree graph with the contents (Fig. 8). Each vertex represents either a director (red square) or a writer (green circle), and the nodes are titles with an average rating above 7, differing in color according to this feature. This graph excludes all isolates and portrays only the maximum crew members that are connected through at least one title. A division is clear among this data: On the left, the directors relate themselves through some titles, and then these also relate to some writers, comprehending titles that include more than one director or writer per title. On the other side (right), a significant part of the directors only relates themselves to one writer, which makes us believe that most of the titles portrayed on this side of the graph only have one director and one writer. Given that most of the connections are blue (average ratings between 7 and 8), we conclude that there is no specific kin between the number of members of a crew and the title’s average rating above 8. Despite this, one specific director (second on the top) is heavily related to many others, which can indicate that making a movie with this director may lead to success.

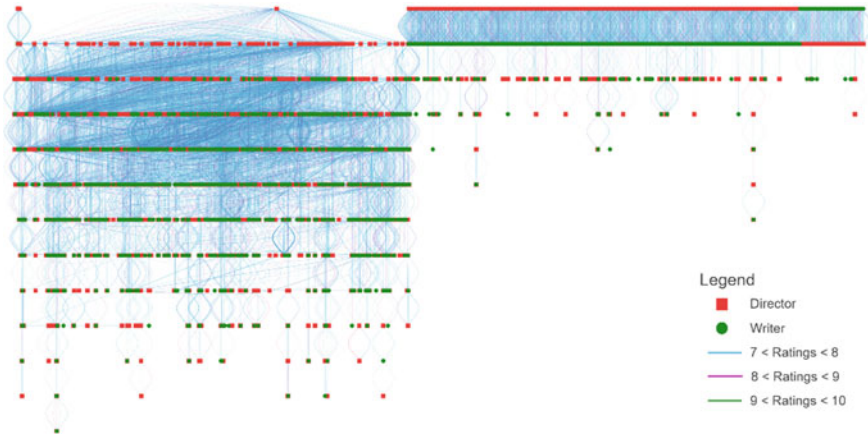


Fig. 8 Tree showing the relationships among the crew of successful titles

Title’s Awards: The *Emmy*, the *Golden Globes*, and the *Oscar* awards datasets have 1649, 3808, and 8997 correspondences with *Title Basics* dataset, respectively. In Golden Globes dataset, 727 correspondences are nominations to titles itself, and 3081 nominations are for people in a certain title.

After this merge, we are able to display the distribution of the average ratings (Fig. 9) showing that titles with no nominations have a distribution of ratings with a much wider range than those which had nominations, where the majority of titles with nominations and wins have their IMDb average ratings above 5. Also, the distribution of titles with nominations and with winnings is very similar, although titles with wins only account for 27% of the titles with nominations. We notice that outliers from the distribution of titles with nominations do not appear as much in the titles with wins.

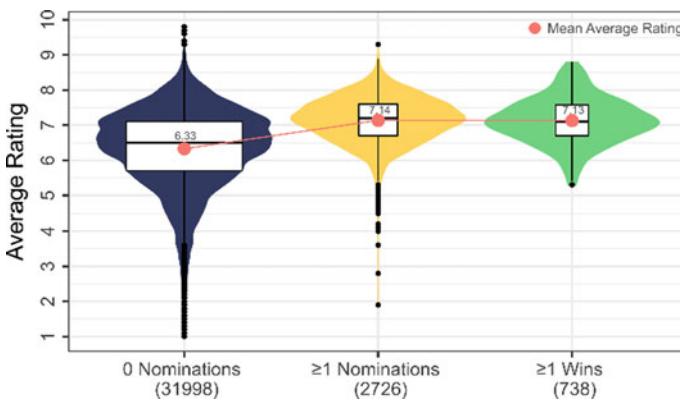


Fig. 9 Distribution of average ratings across titles

Considering only the titles that won an award, we also inspected how their genres may vary. Hence, we plotted the number of titles with each genre (Fig. 10, left) and concluded that drama is the most popular genre for titles that won any type of award studied here, followed by romance, comedy, and biography. We separated these features into titles with 1, 2, or 3 genres, resulting in the middle graphic from Fig. 10. This division only affects the distribution inside each bin. From this, we can conclude that drama appears mostly in titles with three genres, but it also has a considerable part in titles with two or even one genres. It is also the most popular genre across titles with only one genre. We also evaluated if the position of the genre (first, second, and third) is an important feature of this analysis in Fig. 10. From this, we can take that most drama titles have this genre as their first genre, but a considerable amount has genre in the second position. Also, most of the biography titles have this genre as their first genre, which indicates that the position of the genre is important for defining the type of the movie. Going further, we depicted the genres for titles (techniques described in [5]) with awards (Fig. 10). Drama is clearly the number one, and the second place is dominated by romance unless for the Emmy winners, where short comes after the drama. Hence, not considering drama, a movie is more likely to get a Golden Globe for the people if it is a romance, and for the movie if it is a comedy. Since we could separate the nominees that were people, we inspected how they relate through the titles they were nominated for.

The resulting graph showing these relations is in Fig. 11, where each vertex is a nominee, and each node is a title. Nominees that did not win an award are more likely to be connected to other nominees that also did not win, regardless of the movie ratings. Some nominees participate in titles with high ratings but never won

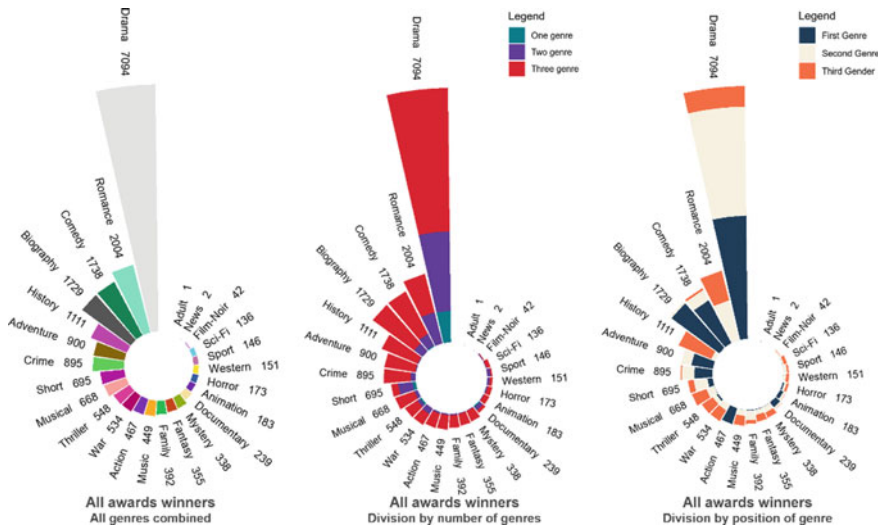


Fig. 10 Genres in titles that won some award

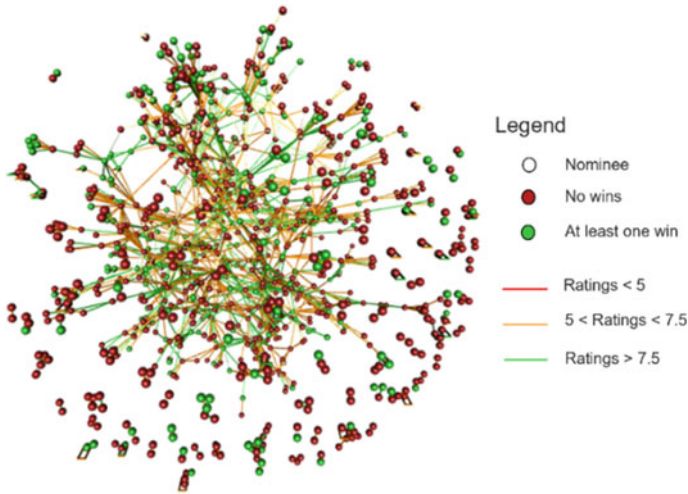


Fig. 11 Relationship between nominees by the nominated titles

an award. Besides this, there is no relevant connection between the rating and the winning.

5 Conclusions

Concerning the regions of titles, the USA is the country to which most titles are translated. Furthermore, titles that are translated to more than one region are typically translated to a European country. As the number of regions increases, the average ratings get higher, and the frequency of translated titles diminishes. Concerning cast and crew, from the analysis performed we concluded that actors are the most relevant people for characterizing the title’s success, that is, the average ratings. Despite this, there may be some specific members of the crew, more exactly directors that can be related to a title’s success.

Titles that win any type of the awards explored (Emmy, Golden Globes, or Oscars) have significantly higher ratings. Drama is clearly the predominant genre for titles that win awards, and genre features may determine the type of nominee of an award.

Taking all of this into consideration, our analysis responds to how can a title be successful according to the studied terms.

Acknowledgements This work is financed by National Funds through the Portuguese funding agency, FCT—Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

References

1. Canet Centellas, F.J., Valero Navarro, M.A., Codina Bonilla, L.: Quantitative approaches for evaluating the influence of films using the IMDb database. *Commun. Soc.*, **29**(2), 151–172 (2016). <https://doi.org/10.15581/003.29.2.151-172>
2. Wilke, C. O. (2019). *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media. ISBN-10:1492031089
3. Sievert, C.: *Interactive web-based data visualization with R, plotly, and shiny*. CRC Press. ISBN-10: 1138331457 (2020)
4. Oghina, A., Breuss, M., Tsagkias, M., De Rijke, M.: Predicting imdb movie ratings using social media. In *European conference on information retrieval* (pp. 503–507). Springer, Berlin, Heidelberg, (Apr 2012)
5. Ward, M. O., Grinstein, G., Keim, D.: *Interactive data visualization: foundations, techniques, and applications*. 2nd Edn. CRC press. ISBN-10: 0367783487 (2021)

Impact of COVID-19 and Rapid Response of Small Restaurants in Thailand



Saruda Sunthornpan and Sadayo Hirata

Abstract The restaurant and service industry is a very important contributor to the global economy and the Thai economy. In particular, small restaurant in Bangkok, the capital city of Thailand, accounts for 41.78% of the total restaurants in the country. However, these sectors are vulnerable and susceptible to the COVID-19 outbreak. This research aims to report the impact of COVID-19 on small restaurants and consider the rapid short-term measures that restaurants are taking in response to the crisis. It is based on the latest fact-finding surveys to provide real-life solutions to problems. This research collected data from 290 respondents from small restaurant business via interview by telephone. We found the level of impact was directly related to sales and staffing rates, including making decisions about the business status as well as the rapid response of restaurants depending on consumer behavior, society, and technology to increase sales and reach consumers in this era.

Keywords COVID-19 · Small restaurant · Response · Technology

1 Introduction

1.1 Importance of Restaurants Businesses

Today, consumers want more convenience and speed. Therefore, they prefer to eat out more. Restaurants play an important role in meeting the needs of consumers. It is one of the businesses that are extremely important to the country's economic sector.

Thailand has a large number of restaurants and is known to both Thais and foreigners. At the beginning of 2020, restaurant business in Thailand has a turnover value of not less than 124 billion US dollar (USD). It accounts for 4.7% of the total service sector of gross domestic product (GDP service sector). In 2020, Kasikorn

S. Sunthornpan (✉) · S. Hirata
Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan
e-mail: na20503@shibaura-it.ac.jp

S. Hirata
e-mail: hirata-s@shibaura-it.ac.jp

Bank Research Center (Thailand) expects the restaurant business to be valued at 135–138 billion USD or expand by 1.4–2.4% from the previous year and has a tendency to increase steadily [1].

The importance of restaurant businesses segment in Thailand is small restaurant because it is located everywhere in Thailand. More than 42.21% of restaurants are located in Bangkok, the capital city [2]. In addition, small restaurants are many of businesses accounting for 22.79% of small business and have high employment rates accounting for 15.37% of total employment rate in Thailand. It can be seen that the restaurant business can generate a large amount of labor and become a source of income for the middle and lower classes of the community [3]. It can be seen small restaurants are very important to Thailand's economy.

1.2 Coronavirus and Impact for Restaurant Businesses

At the end of 2019, there was a major outbreak of coronavirus disease (COVID-19) around the world. It not only affects health but also the economy especially restaurant businesses. The government has recommended or directed restaurant businesses to focus on delivery services or reduce the number of seats due to social distancing policies. The forecast for the restaurant's future is disastrous. Experts estimate that more than half of restaurants will not survive [4]. Additionally, several governments in Europe, America, and East Asia, including Thailand, have temporarily closed bars and restaurants as a final measure to stop the epidemic [5].

According to the National Restaurant Association survey of more than 6000 restaurant owners. Income in mid-April About 60% of US restaurants have been forced to close due to financial hardship [6].

The Indian restaurant industry is experiencing a 50–70% drop in revenue this year due to disruptions caused by the pandemic. Dine-in restaurants now rely on home delivery and takeout services. It also leads to restaurant closures and employee loss [7].

In the case of a Portuguese restaurant, it was found that 75% of restaurants were close for two months and only 25% were open by takeout and/or delivery. It also stated that more than 70% of restaurants would not be able to pay employee's salaries without timely government support. More than 30% of restaurants are considering bankruptcy [8]. The epidemic continues, challenges will increase in terms of income generation, employment, business growth, and the ability to turn capital. And if the spread of the COVID-19 virus increases, the proportion of restaurants reopening or starting a new business will be greatly reduced [9].

Thailand started to have infected in early 2020. During COVID-19 in 2020, small restaurants business in Thailand contracted by 5.8%. Thailand has many challenges due to the negligent lifestyle of people who do not follow the rules—social distancing leads to rapid spread of disease, disease management measures and delayed procurement of vaccines including the recession in the country. Now, the number of infected

patients continues to increase. The Office of Small and Medium Enterprises Promotion (OSMEP) reported in March 2021, small restaurants business closed business permanently 5.3% and propose a small business adaptation plan during COVID-19 in the country which does not have a recommendation that focuses on a particular business [2]. We know the state of the business is deteriorating. But we do not know the real problem. We therefore do research to understand the real problems and understand rapid response measures of restaurant business.

This research on COVID-19 is important and aims to report the impact of COVID-19 on small restaurants and consider the rapid short-term measures that restaurants are taking in response to the crisis. It is based on the latest fact-finding surveys to provide real-life solutions to problems.

The researchers believe that if we know the severity of the impact, anxiety, and decision-making within the business. It will help policymakers to plan, adopt policies or regulations to the needs of small restaurant operators, which are the core of the restaurant business. This will reinforce the economic confidence of the restaurant business in the future.

2 Methodology

We collected data from small restaurant in Bangkok. The data were collected by telephone using a questionnaire. This method suitable to collect information comprehensively and quickly also saves money compared with going out to collect data [8].

We collected data from February 22, 2021, to June 01, 2021, from 290 respondents from restaurant business in Bangkok, Thailand, that approximately 85% of the respondents were business owners or managers. They had the authority to make decisions for the business. Business data were analyzed using descriptive statistics.

The questionnaire was evaluated by two experts and three restaurant owners. The questionnaire asks about the level of impact of COVID-19 and other dimensions that have changed from the impact of COVID-19 such as sales, headcount, business status including rapid response crisis. Our research principles are to inform the research objectives in full, obtain consent, maintain confidentiality, and respect respondents [10] (Table 1).

3 Result

3.1 Impact of COVID-19 on Small Restaurants in Thailand

COVID-19 has affected small restaurant businesses in Bangkok on several levels: the majority of restaurant businesses (62.07%) suffered a very serious impact, representing 71–100% severity, followed by 25.17% of small restaurant that suffered a largely negative impact (51–75% severity), 11.38% of small businesses that suffered a small negative impact (26–50% severity), and only 1.38% of businesses that suffered a slightly negative impact (1–25% severity). The results are presented in Fig. 1.

Smaller restaurants affected by a slightly negative impact saw a 25% increase in sales, a 25% sales remain the same (no change), and the majority of smaller restaurants affected by a slightly negative impact accounting for 50% drop in sales (21% = <40%). The majority of smaller restaurants affected by a small negative impact saw a 42.42% drop in sales (41% = <60%). Impact of largely negative affected restaurants accounted for 63% found that sales decreased (61% = <80%), which is considered a significant reduction in sales began to affect business operations. The majority of smaller restaurants affected by a very serious impact accounting for 66.11% drop in sales (81% = <100%). The results are presented in Fig. 2.

Studies have shown that the degree of COVID impact is directly related to sales. It can be seen that the restaurant was slightly affected, sales have hardly changed. Compared to restaurants that were very seriously affected, sales have dropped dramatically.

Almost all the declining sales resulted in the need for operators to find urgent solutions and some decisions to keep their stores alive, which we will discuss in the next section of results.

We surveyed about the number of full-time employees in a small restaurant. It was found that the rate of staff reductions was also correlated with the level of impact caused by COVID-19.

Restaurants affected by the slight negative impact still retained the same staff as the most, at 75% and the proportion of staff reductions only 25%. Small negative

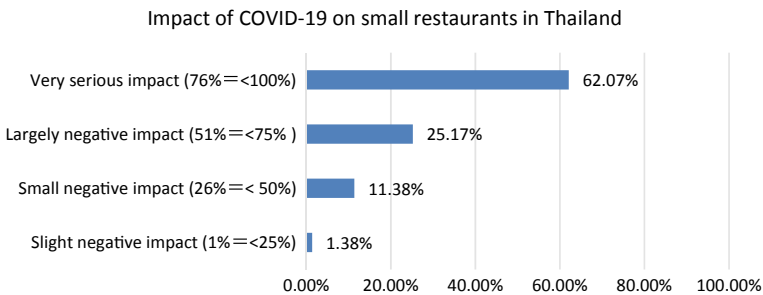


Fig. 1 Level impact of COVID-19 on small restaurants in Thailand

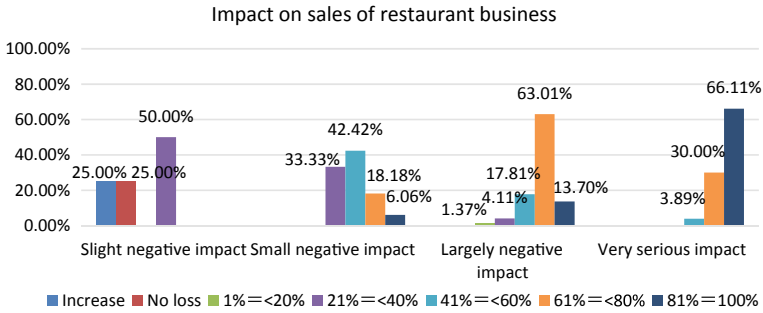


Fig. 2 Correlation of the degree of impact of COVID and sales of small restaurant business

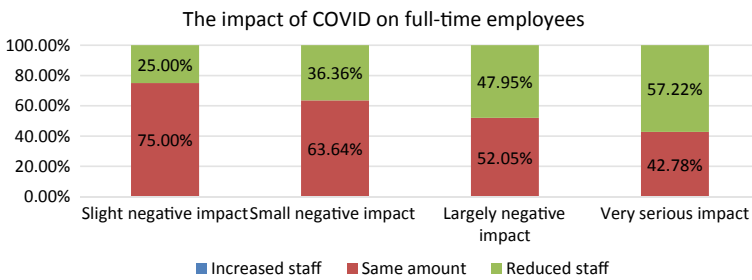


Fig. 3 Impact of COVID on full-time employees in restaurant business

impact restaurants still retained the same staff at 63.64% and a staff reduction rate of 36.36%. Stores affected by largely negative impact still retained the same staff at 52.05% and a staff reduction rate of 47.95%. Stores affected by very serious impact found that still retained the same staff at 42.78% and a staff reduction rate up to 57.22%. The results are presented in Fig. 3.

Most restaurants rarely lay off full-time employees. Due to the termination of employment will have to pay a large amount of compensation. Most restaurants have to be repressed to hire staff. They may reduce working days or pay less wages to survive.

In addition to the full-time employees affected, daily workers are also affected. Daily employees including part-time job are important to the restaurant business as well, that employ daily workers to taking orders, serving or even cashier. Restaurants are also source of income for daily employees such as students and people in the community.

We will find that the layoff rate of daily employees will increase with the effect of the restaurant received from the COVID-19 crisis. For example, we did not find any daily employee layoffs in restaurants that received a slight negative level. On the contrary, we have found that very serious affected restaurants decided to cut their daily workforce by as much as 82.52% in order to cut costs. Because a large percentage of the cost in the store is staff wages, the results are shown in Fig. 4.

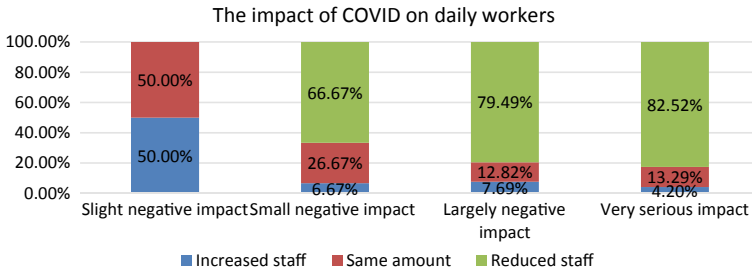


Fig. 4 Impact of COVID on daily worker in restaurant business

We also saw an increase in the number of daily employees being hired. Part of the layoff of full-time employees/full-time employees resigning and hiring daily workers to replace them. And there are some restaurants that have changed their full-time staff to be daily employees instead. It is a voluntary change of permanent employees. This way, the restaurant’s wages will be reduced (paid according to working hours) and cut down benefits as well to support the restaurant to survive in this difficult situation.

From our research on the impact of COVID-19, we have found that the slightly negative affected restaurants were still 100% operational. Restaurants affected by small negative impact found that 87.88% were open and decided to reduce the number of branches by 12.12%. Followed by restaurants that had a large negative impact, with 72.60% still operating, 24.66% reduce the number of branches and 2.74% closing down temporarily and restaurants affected by a very serious impact, 67.22% remain open, 16.11% decide to cut down, 7.78% decide to close temporarily and 8.89% decide to close permanently. Due to a large decrease in sale including the expenses that entrepreneurs are responsible for the month is so high. Some operators need to find a way to respond to this dire situation until they decide to close the shop permanently. The results as shown in Fig. 5.

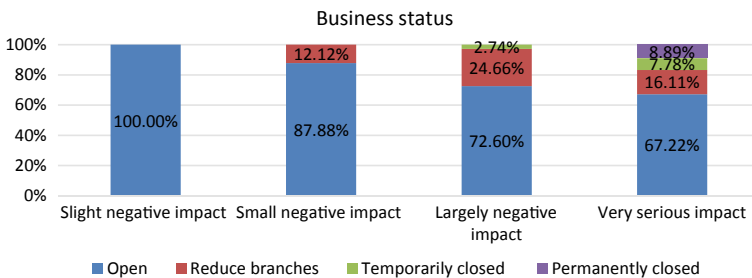


Fig. 5 Business status of restaurant in each level of impact

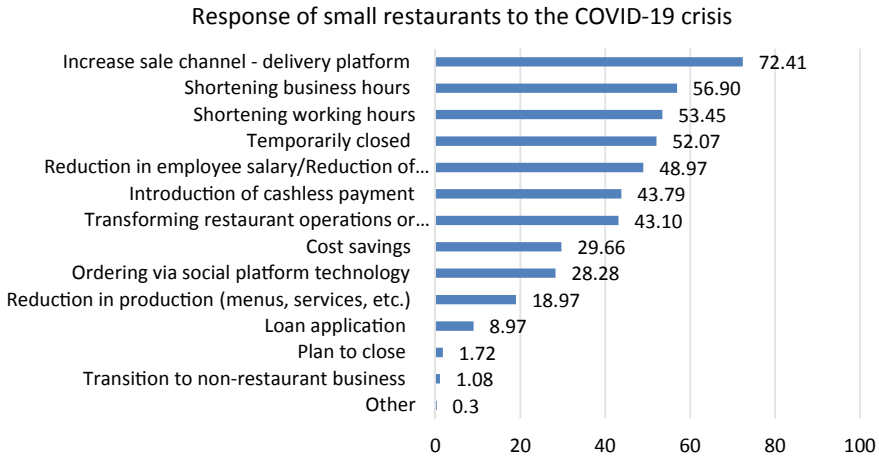


Fig. 6 Response of small restaurants to the COVID-19 crisis

3.2 Rapid Response of Small Restaurants to the COVID-19

The study found that small restaurants. There are several ways to respond quickly to COVID-19. It was found that more than 70% of small restaurants chose an increase sale channel—delivery platform. It can be seen that in this era, technology is important to reach consumers. Followed by shortening working hours 56.90% and shortening business hours 53.45%. Following by temporarily closed 52.07%, reduction in employee salary/reduction of employees 48.97%, introduction of cashless payment 43.97%, transforming restaurant operations or business models 43.10%, cost savings 29.66%, ordering via social platform technology 28.28%, reduction in production (menus, services, etc.) 18.97%, loan application 8.97%, plan to close 1.72% and transition to non-restaurant business 1.38%. The results as shown in Fig. 6.

An in-depth analysis reveals that small restaurants affected in each level have different response decisions as shown in Fig. 7. Small restaurants affected by the slight negative impact will respond to COVID-19 with increasing sales channel—delivery platforms and shortening business hours mainly. Small negative impact restaurants will respond by ordering via social platform technology and transforming restaurant operations or business models mainly. Followed by a largely negative impact of smaller restaurants, which would respond by planning to close because he began to see problems that were difficult to manage. The accumulated funds began to run out including not sure when this crisis will pass and some decided to close the shop permanently. Small restaurants affected by a very serious impact also faced financial difficulties due to declining sales. Some people decide to apply for a loan. Some of the menus or services are reduced to save costs. Some have chosen to close their stores permanently and turn to other businesses that are not restaurants. This is a crisis where restaurants are struggling and respond quickly to survive. But if the COVID-19 situation is long, there may be more restaurants to close.

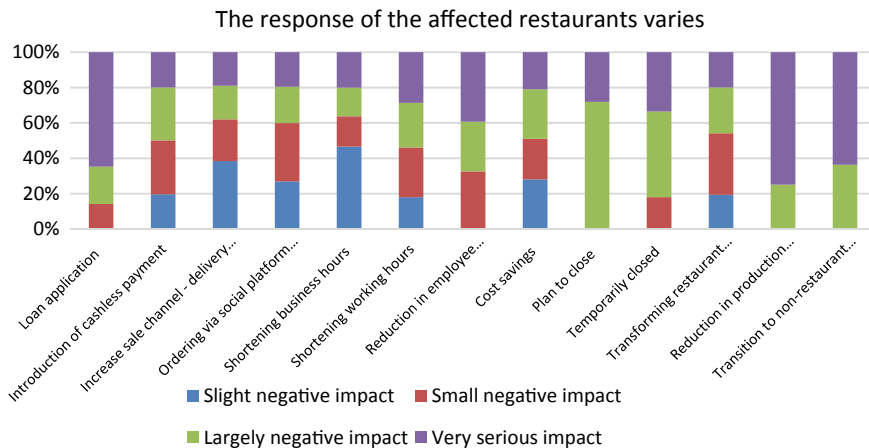


Fig. 7 Response of the restaurants has been affected at different levels

4 Recommendation and Suggestion

The COVID-19 crisis has caused a lot of damage to small restaurants. It can be seen that many small restaurants are severely affected by the decline in sales. Including layoffs of employees. That means economic losses that occur not only in Thailand but all over the world.

Various Indian restaurants have seen a huge drop in consumers. As the trend of eating at home increases, amid pandemic fears. There are experiencing a 50–70% drop in revenue in 2020. There find opportunities to introduce innovations including creating innovations to provide a safe dining experience for customers, for example, contactless experience, digital payments, cloud kitchens, and immunity-boosting food options [7].

COVID-19 has dramatically reduced restaurant sales. Improving the customer experience by promoting services and offering attractive restaurant menus (via website or online platform) can be a strategy to drive sales and generate new revenue. Most of the large chain restaurants have provided an online platform that works well with attractive menus. However, many smaller restaurants do not provide such a system [11]. The way to make successful small restaurant with increase sales through online platforms, them found that consumers’ demand for food and convenient services along with eye-catching pictures and menu. This will give consumers more opportunities to buy food in restaurants [12].

This study addresses one side of the pandemic from a market perspective. Emphasis only on the ability of experts to respond. Regardless of the complexity of other important factors such as policies that support restaurants and technology, readiness of entrepreneurs to new technology, culture, and society are all involved in the survival of the restaurant. That is the weak point of this research, we will study in the future.

5 Conclusion

Our research found that smaller restaurants were affected by a very serious impact 62.07%, followed by largely negative impact 25.17%, small negative impact 11.38%, and slight negative impact only 1.38%. We found that the severity of the impact was directly proportional to sales and employment. This means that small restaurants with a very serious impact (80%-100% decrease in sales) are laying off employees of both full-time and daily employees, including the opportunity to close the shop permanently. However, restaurants have responded quickly to the crisis. More than 70% of small restaurants chose an increase sale channel—delivery platform, followed by shortening working hours, shortening business hours, temporarily closed, reduction in employee salary/reduction of employees and other ways. It can be seen that technology has started to play an increasingly important role in the restaurant business. The researchers hope that this research will help policy planners understand the problems of small restaurant and make measures that are the interests of restaurants in this crisis.

Table 1 Questionnaire design

Questionnaires	
Question 1: Level of the impact of COVID-19 (Overview)	
(One choice)	<ul style="list-style-type: none"> • Positive impact • No impact • Slight negative impact (1% = <25%) • Small negative impact (26% = <50%) • Large negative impact (51% = <75%) • Very serious impact (76% = <100%)
Question 2: Declining sales in 2020	
(One choice)	<ul style="list-style-type: none"> • Increase • No loss • 1% = <20% • 21% = <40% • 41% = <60% • 61% = <80% • 81% = <100%
Question 3: Amount of full-time employee	
(One choice)	<ul style="list-style-type: none"> • Increased staff • Same amount • Reduced staff
Question 4: Amount of part-time employee	
(One choice)	<ul style="list-style-type: none"> • Increased staff • Same amount • Reduced staff

(continued)

Table 1 (continued)

Questionnaires	
Question 1: Level of the impact of COVID-19 (Overview)	
(One choice)	<ul style="list-style-type: none"> • Positive impact • No impact • Slight negative impact (1% = <25%) • Small negative impact (26% = <50%) • Large negative impact (51% = <75%) • Very serious impact (76% = <100%)
Question 5: Business status	
(One choice)	<ul style="list-style-type: none"> • Open • Reduce branches • Temporarily closed • Permanently closed
Question 6: Rapid measures to respond to COVID-19	
(Multiple choice)	<ul style="list-style-type: none"> • Loan application • Introduction of cashless payment • Increase sale channel—delivery platform • Ordering via social platform technology • Shortening business hours • Shortening working hours • Reduction in employee salary/reduction of employees • Cost savings • Plan to close • Temporarily closed • Transforming restaurant operations or business models • Reduction in production (menus, services, etc.) • Transition to non-restaurant business • Other

References

1. Kasikorn Research Homepage, <https://www.kasikornresearch.com/th/analysis/k-econ/business/Pages/z3067.aspx>, Last accessed 17 Sep 2021
2. The Office of Small and Medium Enterprises Promotion (OSMEP), Dashboard SME big data, <https://sme.go.th/th/page.php?modulekey=348>, Last accessed 17 Sep 2021
3. The Office of Small and Medium Enterprises Promotion (OSMEP).: White paper on MSME 2020,” Research and Risk-Warning Analysis Division, Department of Industrial Promotion, Ministry of Industry, Thailand, (2020)
4. Jungkeun, K., Jacob, C.L.: Effects of COVID-19 on preferences for private dining facilities in restaurants. *J. Hosp. Tour. Manag.* **45**, 67–70 (2020)
5. Aristeidis, G.: Developing an indicative model for preserving restaurant viability during the COVID-19 crisis. *Tourism Hospital. Res.*, 0(0), 1–14 (2021)
6. Yang, Y., Hongbo, L., Xiang, C.: COVID-19 and restaurant demand: early effects of the pandemic and stay-at-home orders. *Int. J. Contemp. Hosp. Manag.* **32**(12), 3809–3834 (2020)
7. Shinu, V., Richa, N.A.: Repercussions of COVID-19 on small restaurant entrepreneurs: the Indian context. *Strateg. Chang.* **30**, 145–152 (2021)

8. Arlindo, M., Teresa, P., Alexandra, S. M.: The Impact of pandemic crisis on the restaurant business: Sustainability (13), 1–13 (2021)
9. Michael, G. B., Robert E. F., Brumby, M. M., Melinda, O. P.: Independent restaurant operator perspectives in the wake of the COVID-19 pandemic. *Int. J. Hospital. Manag.*, **93**, –8 (2021)
10. De, L.: Data Quality in Mail, Telephone and Face to Face Surveys. TT-Publikaties, Vrije Universiteit Amsterdam (1992)
11. Kate, K., Belinda. C., Vivienne. B., John. S.: Good practice in the conduct and reporting of survey research. *Int. J. Qual. Health Care* **15**, 261–266 (2003)
12. Prawannarat, B., Angela, G.S.: The effect of online restaurant menus on consumers' purchase intentions during the COVID-19 pandemic. *Int. J. Hosp. Manag.* **94**, 1–9 (2021)

Improving Communication with Media: Portuguese National Public Police Case



Sónia M. A. Morgado 

Abstract Good communication between media and National Public Police is one imperative contrivance in thought-provoking in consumers. Yet, communicating processes to date have been conditioned by the unique phenomenon associated with the type of information provided by police. Problems in communicating involve unbalanced, disruptive, unproductively, and disenchantment opinion of police activity. By probing the external communication that bond depictions of media support to consumers' reactions, this research addresses the elements which fundament an improved communication plan to incorporate the legitimacy and fairness of the communication, reliability, and the representation of media, to support the adjustment of the conception of police mission and values in the media and social vicinity. Using the Delphi method, the authors built a model in which an alliance and a crossover effect between past and present fundament the creation of gate-keeping narratives, information messages, dissemination accurateness, and educational proactive approaches, structuring a co-creative environment. The research offers a substantial contribution to the reflection of the managers and may be a powerful element of guidance to enhance the success of interaction between the main stakeholders. This approach develops a model that sustains an idiosyncratic communication framework for determining the content, form, style, and sequence of transmission and reception.

Keywords Communication plan · Disruption · Communication fairness · GIDE model · Legitimacy · Media

S. M. A. Morgado (✉)

ICPOL—Centro de Investigação do Instituto Superior de Ciências Policiais e Segurança Interna,
Lisbon, Portugal

e-mail: smmorgado@psp.pt

CIEQV—Life Quality Research Center, Leiria, Portugal

1 Introduction

Communication in the public sphere must endure rigorous rationalization principles, optimization, and management to avoid misunderstandings, disproportionate, and the unexpected. In this sense, the revision of communication politics is necessary to enable the consideration of the needs of the institutions and target audience. Meirinhos [1] considers that a strategic plan from the organization provides a set of directional references of activity, while the communication plan stipulates the sense and form of its presentation.

The importance of communication is not static entails an evolutive process, which institutions rely on because its development is reinforced by the confidence of receptors, compelling a continuous dialog between [2]. Emery et al. [3] underpin this concept. They confirm that the reach of audiences is only possible by using communications instruments such as radio, television, journals, digital platforms, and others. In this sense, media are active participants in the institution's image construction, help to form ideas, and generate values [4] about a brand, organization, or the imaging of an institution. So, media have a determinant role in the process of cultural conservation and society transformation [5], contribute to news notoriety, and incite higher levels of consumption [6]. On the other hand, media, and even more social media, are the "ideal medium (...) to expose views and reach out" to society, which upheaval's "media populism and online disinformation" (p. 181) [7] and misinterpretation.

Consequently, the National Public Police (Polícia de Segurança Pública—PSP) role of public service is promoting the participation and involvement of all participants. These affect the interest and collective lives [8, 9].

In the current environment, solid communication processes are a requirement, even though is a modest endeavor in public services [10]. By understanding the needs of the society, the institutional objectives are ensuring communication to have a philosophy and strategy of integration of those aspects [11, 12].

To PSP, the communication is paramount to avoid disturbance in the public perception. The brawl may be a result of the regular interaction between police and citizens. It is up to the institution to communicate daily activities as a way to reflect its identity.

This paper incorporates the external communication of the police, to be able to characterize it, identifying its objectives, and explain how to improve the communication process.

2 Conceptual Background

The word origins are embedded in the Latin word *communicationis*, refereeing that if there is no agreement in the process, there is no communication [13]. By this fact, communication plays a meaningful role in society because it is not possible to live without communicating or relating [14]. As a complex process of information production [15], it has a persuasive dimension as stated by Aristoteles [16].

The interactivity of the information allows the development of society, from the transmission of well-founded and essential knowledge for the continuous growth of citizens, organizations, and society. The responsibility of good and effective communication relies upon the information producers rather than on the receptor [17, 18]. A process of communication without disturbances advocated by Shannon and Weaver [19] is utopic because the channels used for the transfer are also its sources and depend not only on the language, quality of the information received but also on the subjectivity and the experience of the receptor.

Some common ground is established considering what makes the communication effective: (i) codification [20, 21]; (ii) change of meanings [22]; (iii) language [23]; (iv) information [24]; (v) human activity [25], and (vi) opportunities creator.

Technology has gained terrain over the years and has become a part of human evolution. In the information society, digital means, communication process [26], and social media [27] communicating have been transferred to a digital environment, which poses transmutations in the process [28, 29].

Organizations and communication are intertwined. The transmission of information can be unilateral [12] but must be public relevant [30] to generate assimilation, interpretation, and learning [31]. Zémor [32] states that there are five main categories for organizational communication: (i) obligation to report to his publics; (ii) creation of the relation with the public; (ii) presentation and promotion of services; (iv) internal and external communication; and (v) development of actions of public interest.

Our focus lies in external communication that is defined as the creation of links between organizations and stakeholders, as to be an integrative part of the organization [33], contributing to form its identity [34] and the image and reputation [35].

Adjustments on the vision and mission statements and the formulation of organizations goals and objectives concerning the communication plan are a strategy to achieve visibility and legitimacy [36]. Strategic communication is persuasive [37] and more efficient than public relations [38]. How media receives information is essential because the correct dissemination allows people to process information and being judgmental with the legitimate arguments, even if case-based, rather than on the whole institutional behavior [39].

The information obtained from the communication of institutions that comply with democratic states endeavor is the main element for combating disruption of communication, the disinformation order, and consequently the potential dwindle of democratic institutions [40].

Establishing a communication plan is the way to align it according to institutional strategy [41]. On the other hand, while guaranteeing the trust in democratic institutions, it avoids the growth of the credulous public, which might affect the quality, the understanding, and the veracity of the information. PSP is not unaware of the benefits of having a plan, and in the major strategic options (GOP) 2017–2020 [42], there was the consideration of the external communication with the following measures: (i) new web page; (ii) PSP APP; (iii) redefinition of communication strategy; and (iv) concretization of the plan of communication. These follow previous orientations

that included the setting up of the Press and Public Relations Department (GIRP), the definition of the stakeholders, and the communication manual. PSP is no exception, and the communication plan (internal and external) is also a part of the management toolkit.

3 Methodology

To investigate the hypothesis, in this exploratory study of qualitative nature, we combined a literature review of the communication in law enforcement with the Delphi technique about the aspects necessary to improve PSP communication. The principal approach is the qualitative method which is prevailing in communication research [43]. Daba-Buzoianu et al. [44] states that “the epistemology of communication is deeply linked with qualitative inquiries” (p. 1). The investigation also considers the quantitative analysis for testing some hypotheses.

3.1 *Conceptual Model and Hypothesis*

Accordingly, with the introduction, an organization must have an effective means of communication.

Our purpose with building a conceptual model is the excavation of the subject, tested by the scientific rigors of logic, facts, and statistical empowerment. In fact, for resolving problems, more than descriptive knowledge [45] is necessary to ask questions, choose variables, describe objectives, and the population in the study [46].

The perception of the logic between propositions that enables the definition of a hypothesis allows deriving a conceptual model (Fig. 1). The literature review and objectives proposed are the basis for the model. It is also the leverage for the upgrading of the communication process. Conceptual models are tools adequate to represent knowledge [47] and for case studies [48].

In this view and because the search of the facts is the main objective of the investigation [49], the hypotheses are assertions that compel to learn and explore the relations between concepts [50, 51].

Benn et al. [52], Boros [53], Brown et al. [54], and Carrilho [11] suggested that organizations must understand the outside stakeholders as elements of the institution because they are targets, but also the receptor of information. Drawing on these conclusions, the public positioning is due to stakeholders, and how they receive and perceive information, a first hypothesis was identified:

H1: The performance of the PSP for external communication is considered positive by the media.

Barichello and Dall’Agnese [36], Cunha et al. [55], França [56], Kent and Taylor [57, 58], and Moreno [59] stated that the leading role of media is the dissemination of information, and organizations should use it to reach out the public. They are



Fig. 1 Conceptual model

leverage in the prosecution of institutional objectives, which leads to the following hypothesis:

H2: The current strategy of the PSP external communication answers the needs of the media.

At last, we raise the relatable standpoint studied by França [56], Kotler and Keller [60], Meirinhos [1], and Smith [37] that the communication is a strategic action. Alongside these authors, Steyn and Puth [41] revealed the connection between planning and effective communication and as a resource and mandate to changing patterns and perceptions [61]. Thus, the following hypothesis was tested:

H3: A communication plan is a strategic option for the PSP as a management instrument.

The present conceptual hypothesis is the mainframe for defining the elements of an upgrade and improved communication plan for PSP.

3.2 Method

Based on the scientific method, within exploratory, descriptive, and correlational approach, the resource for determining the level of interlinked and consensus between questions was the Delphi method [62].

The technique was applied to a panel of experts that comprises the media mailing list of the PSP (journalists, pressman, newspapers editors, magazine publishers, of social media traditional, online mass media). The goal was to achieve convergence concerning the communication plan. Even though the core mission is not at the communication level, an external communication plan is a strategic management tool and must be planned, assessed, and efficient to evolve to an accurate public perception of the police activities.

Considering the need of stimulating the development of new ideas [63], the lack of information about the substance of matter [62, 63], the dimension of the mailing list, and the heterogeneity of backgrounds, the technique is adjusted to study [64], for consensus building.

This technique is essential to structure a group communication because it allows dealing with a complex problem [65] and withstands the anonymity of the participants, the representation of the results, and the feedback to the response for a re-evaluation of the results [66].

In deference to qualitative methodology, this technique is reliable [67] and effective because the knowledge and experience of the experts sustain it [63].

Descriptive measures demonstrate and reveal the search for consensus [68–71].

On the other hand, a Chi-square measure was applied, to perform the examination of the differences between categorial variables according to analyze the performance of PSP as best actor in the relation with the media.

The Chi-square test is a statistical procedure used by researchers to examine the differences between categorial variables in the same population.

3.3 Corpus and Procedures

The data were recollected by a questionnaire sent to the panel comprised of experts relating to media. The professionals, media experts, in the mails list of GIRP constitute a dynamic panel with professional experience in divergent fields of expertise.

There was a total of 28 informants, composed of 53.6% of females and 46.4% of males, ranging and an average age between 31 and 40 years old. The majority is single (50%) and had a university degree (92.9%).

Concerning the media, the press and television represented 42.9% each and the radio 14.3%. The analysis considered traditional media. Fifty-seven percent of the experts had more than 15 years of experience.

Following the study objectives, the panel comprehended of media experts received the questionnaire. They are the external target audience and a mediator between PSP and the general public. The questionnaire was available to the experts by a hyperlink on Google docs. It was in a structured format which allowed the extraction of data for future analysis. The respondents were encouraged to participate and asked to grade the statements regarding an inverted Likert scale (5—completely disagree, 1—completely agree).

Basing the threshold for judging the level of agreement among the experts when the scoring was equal or lower than 3 and if 50% of the topics received the majority of votes [71].

The first section of the questionnaire included the objectives and consent form and socio-demographic data, followed by the questions concerning communication.

The data were analyzed using Statistical Program for Social Sciences (SPSS, Inc. EUA), version 24, and Excel 2007 (Microsoft Corporation). The level of confidence was $p \leq 0.05$.

4 Results

Through the exploration of the results, the first step was considering the questions according to the dimensions: 1—the importance of communication activity for PSP; 2—perspective of the media as beneficiaries of PSP communication; 3—characteristics of the communication actions from the media; 4—relevant aspects of institutional communication; 5—pertinence of the current PSP initiatives in promoting institutional communication in the media.

On the questions relating to the first dimension, there was a unanimous consensus about how meaningful communication for the PSP is, with 85.7% of the experts completely agreeing about the importance. About 42.9% of the experts also revealed satisfaction with the current communication strategy, which validates its performance. The turning point of the positive approach is that most experts express uncertainty about exemplary behavior in the communication process. One may assert that this uncertainty is not evanescent but stranded in the hollowing of the need for communication consumption, without the veil of confidentiality, which sometimes is associated with police activity, to maintain public order.

For this aspect, the experts consider that the output of PSP external communication is valued and accepted by the media as adequate. Bearing in mind that the agreement of the results is not successful in every question, the first hypothesis is not confirmed. Yet, despite the results, we cannot overcome the “relationship between press visibility and public knowledge” (p. 16) [6] so relevant in the context of PSP communication.

As beneficiaries and users of the external communication, the experts reported that they frequently received news from PSP. It is possible because the consensus resulted from applying cumulative criteria (median and means). The agreement is also acquired relating to a lack of interest in the Web site of PSP. As so, the utility of the page to the media is scarce. This disinterest also covers the YouTube channel, with a consensual agreement. In other areas, there was less consensus concerning Facebook, Twitter, or Instagram. Given the results, the media are not considering social media as a relevant aspect of the communication process.

Regarding the characteristics, the discourse meets the media needs and is objective (mean = 1.25; 1.29; median = 1; 1, respectively); even though there is a lack of common ground concerning the existence or non-existence of delays, the timing of news to the newsroom. It might be due to the tight schedule of the newsrooms. The factor time is incompatible with the police’s need for verifying facts before the press release. It is also necessary to mention that PSP is the primary news source concerning operational activity as PSP (mean = 1.32, median = 1).

Even in the borderline of having consensus, the fact is that only five of the six questions have unison. Hypothesis two is not confirmed.

The non-confirmation of H1 and H2 reveals that the police is not adjusting the communication to the media. If it were, the strategy was optimal, as stated by Barichello and Dall'Agnese [36], and França [56].

From the analysis of the fourth dimension, it is clear that PSP should: (i) promote protocols (75%); (ii) interviews and reports; (iii) do more press releases (78.6%); (iv) reinforce the social media; (v) create a platform with operational data accessible to the media (60.7%); and (vi) promote regular meetings with editor-in-chief (78.6%).

It seems that media is eager to be in the spectrum of accessing data promptly by making protocols to improve communicational relations, maintaining press releases. On the other hand, the development of a structured platform according to the police objectives is beheld as a tool of public interest with information available according to the media's needs. Promotion meetings should change the non-consensus answers concerning the quality of information, accessing the digital toolkits (web page, YouTube, social media).

The respondents felt that PSP initiatives, such as expositions (89.2%), events (78.6%), conferences (89.3%), demonstrations (78.6%), publications (75%), partnerships (64.3%), interviews (96.4%), are an important mean to consolidate the communication strategy.

Summing it up, the respondents agree with the importance of communication with quality from the PSP concerning the operational activity. It is a way of interplaying the growing complexity of media systems, combating the disruption, populism, and radical interpretations, avoiding entailing dubious information [73, 74]. That additionally eludes the awareness of knowledge gaps that can provide an "unequal use of media information even if the available information diet is the same" (p. 72) [75].

Testing if the choice as the best force security in communicating has its tole in the satisfaction for the current of the PSP communication conveys the end of the analysis.

Considering the results in Table 1 is possible to determine that the experts scrutinize the PSP as the best communicator among the GNR and PJ (p -value = $0.02 \leq 0.05$). The recognition of good communication is entwined with: respecting the media language, criteria objectivity, and the importance of the information.

To withstand the communication strategy experts, refer to tools such as clipping, email, press releases, announcements, bulletins, press conferences. The non-existence of a structured external communication plan in PSP is a deterrent for a more prophylactic approach and understanding of the police activity. According to Meirinhos [1], Smith [37], and Steyn and Puth [41] pro-organizational purposes, the elaboration of a communication plan is a mandatory tool for PSP. PSP [42] considers the concretization of this kind of plan in the GOP's envisaged for the triennium 2017–2020. In this sense, it is possible to confirm the third hypothesis.

Efforts to introduce a communication plan model in police are advisable [37, 41] and may have been composed as the model GIDE (Fig. 2) that emerged from the aspects recollected:

Table 1 Testing the relevance of PSP as the law enforcement best communicator

Question	χ^2	Df	sig
Satisfaction with present communication	11,643	4	0.020*
Respects media language	10	3	0.024*
Objective criteria	37,714	3	0.000*
Right timing for the newsroom	5429	3	0.143
Communication with necessary information	11,643	4	0.020*
Knowledge of operational activity	3714	3	0.294
Expositions	8429	4	0.024*
Events	23,429	4	0.000*
Conferences	26,000	3	0.000*
Demonstrations	18,876	4	0.001*
Publications	16,286	3	0.001*
Partnerships	8000	3	0.046*
Interviews	15,500	2	0.000*

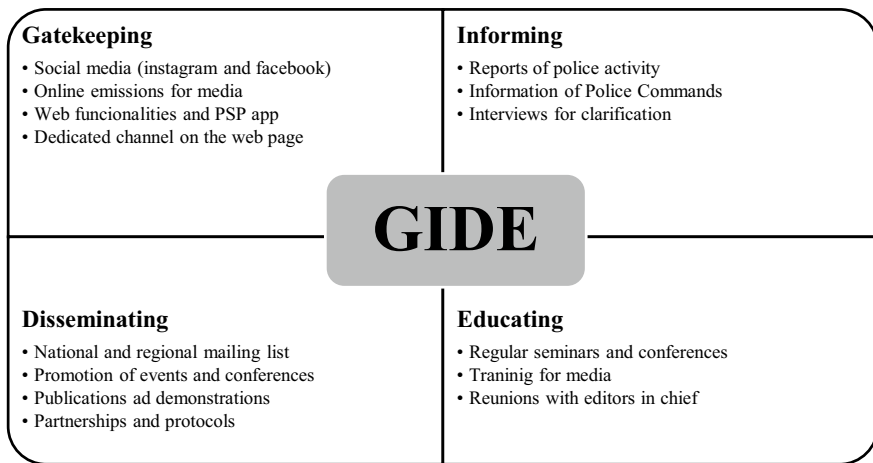


Fig. 2 Communication model GIDE for improving communication

To bring the expert vision to an empirical approach, GIDE model is presented. The locus of the model is the details provided by the specialists such as: to make some investments (e.g., mailing list, reports of police activity, interviews), to maintain the strategies that work (for instance, social media of rapid and visual stimuli like Instagram and Facebook, online emissions for media), and to drop out some strategies that are not so meaningful concerning communication with the media by the police (Twitter, random talk shows, among others) (Table 2).

Table 2 Proposal for an external communication plan for media

Maintaining	Investment	Disinvestment
Instagram (photographs and publications)	National and regional mailing list	Twitter for rapid response
Facebook (photographs and publications)	Dedicated channel on the web page	YouTube for promotional videos
Daily update of social media	Reports of police activity	Random talk shows
Online emissions for media	Interviews for clarifications	Press releases
Information of police commands	Regular seminars and conferences	Email correspondence
Promotion of events and conferences	Training for media	
Publications and demonstrations	Reunions with editors in chief	
Partnerships and protocols	Web functionalities and PSP APP	

Adapted from “Communication planning: A template for organizational change”, by Newman [76]

In effect, the model removes some tools that considered less effective and suggests the introduction of others to encourage police to act competitively to improve the overall communication. Thus, police compete for a higher positive perception of the audiences through the mediators and first-line receivers of information, the media. The media are the mainstream of rampant cultural processes, perceptions and stimulate interest in many areas of public affairs [6], including police activity. This conviction is the core to draw up a plan of communication [1]. It bypasses the recourse to untrustworthy measures that uncover the resources’ inefficiency [11], knowledge gaps [75], and misinterpretation of information.

The media are the first beneficiaries of the need of widespread the message, being the receptors but also the mediators between police and society [36], avoiding the disruption lumped by populism that overcomes above institutional order in a cantankerous public [40], or harshness of public voices [77], that are exhaustive consumers of the headlights.

5 Conclusion

The results of this study corroborate previous research findings shedding light and providing the understanding needed concerning the communication process and plan of PSP.

The comprehensive approach to plan communication is a subject of protracted discussion. From a systemic point of view, human communication is inherited. Globalization, technological evolution, the importance of being informed, the fluxes of information in society cover the fundamental need for collective and individual survival.

Society's perception of police is fundamental. In the current social structure, where media seek to attract more audience, the capacity for reporting unadulterated information can be propitious to generate positive perceptions concerning police activity. For this reason, it is crucial to review and configure the external communication of PSP. It empowers PSP to deal with media situations and biased collective opinion.

Decision-making supported by a communication plan is a mixture of public order, national security, and political approaches. It seems plausible to assume that the decision-making process is constrained by security issues rather than by information or collective needs. However, the media might argue that their need for information is higher and guided by the interest of the audience and their level of news consumption. The difficulty, however, is that they often lack understanding of the impossibility to disclose sensitive police information. Feeding information to media must be a priority in the strategic public exposure of PSP, to undermine the gaps in interpretation, to gatekeep the truth, to provide a better understanding of the use proportionate and adequate use of resources according to the circumstances. Police should communicate to promote comprehension and legitimacy of the actions by its stakeholders.

Inherently, the nature of public service needs the projection of the institutional message. As first responders of police news consumption, receptor, and mediators, the media have the implicit role in restraining disruption, exacerbation of misinterpretation of the public, the ruggedness, and belligerency of the public comments that are proactive consumers of images, videos and newspapers, and social media mastheads. For this role to be promptly and active, the GIDE model is the tool capable of providing this vicinity.

Admittedly, the relation between media and police endures an akin path. Good relations sustained by effective communication can facilitate a proactive rather than a reactive response to facts.

The external communication plan presented is the core element for building the strategy to keep the audiences informed and satisfied. It is also a way of diverting disruptions from media scrutiny by inverting it and deflecting moral judgments.

References

1. Meirinhos, G.D.: Plano de comunicação Instrumento estratégico de gestão. Universidade de Trás-os-Montes e Alto Douro, Minho (2015)
2. Caetano, J., Rasquilha, L.: *Gestão da comunicação*. Quimera, Forte da Casa (2004)
3. Emery, E., Ault, P. H., Agee, W.K.: *Introduction to Mass Communications*. Dodd, Mead & Company, New York City (1971)
4. Pais, L.G., Felgueiras, S., Rodrigues, A., Santos, J., Varela, T.: Protesto político e atividade policial: A perceção dos media. *Revista do Instituto de Ciências Sociais da Universidade de Lisboa* **3**(216), 494–517 (2015)
5. Giacomantonio, M.: *Os meios audiovisuais*. Edições 70, Lisboa (1970)
6. Curran, J., Iyengar, S., Brink Lund, A., Salovaara-Moring, I.: Media system, public knowledge and democracy: a comparative study. *Eur. J. Commun.* **24**(1), 5–26 (2009)
7. Corbu, N., Negrea-Busuioac, E.: Populism meets fake news: social media, stereotypes and emotions. In Krämer, B., Holtz-Bacha, C. (eds.) *Perspectives on Populism and the Media: Avenues for Research*, vol. 7, 1st edn., pp. 181–199. Nomos Verlagsgesellschaft, Baden-Baden (2020)
8. Jeudy, H.-P.: *A Sociedade transbordante*. Edições Sécuro XXI, Porto (1995)
9. Paulino, F.O., Oliveira, L.G.: Serviço público de média e comunicação pública: Conceito, contextos e experiências. *Comunicação e Sociedade* **30**, 55–70 (2016)
10. Stern, L.A., Taylor, K.: Social networking on Facebook. *J. Commun. Speech Theatre Assoc. North Dakota* **20**, 9–20 (2007)
11. Carrillo, M.V.: Comunicação estratégica no ambiente comunicativo das organizações atuais. *Comunicação e Sociedade* **26**, 71–80 (2014)
12. Kunsch, M.M.: Prefácio. In: Porém, M.E., Hidalgo, J., Yaguache, J. (eds.) *Inovações em Relações Públicas e Comunicação Estratégica*, pp. 8–9. Ria Editorial, Aveiro (2019)
13. Lunenburg, F.C.: Communication: the process, barriers, and improving effectiveness. *Schooling* **1**(1), 1–11 (2010)
14. Rüdiger, F.: *As teorias da comunicação*. Artmed Editora, Porto Alegre (2011)
15. Losee, R.M.: Communication defined as complementary informative processes. *J. Inf. Commun. Libr. Sci.* **5**(3), 1–15 (1999)
16. Rosenfield, L.W.: *Aristotle and Information Theory*. De Gruyter, Berlin (2011)
17. Lidstone, J.: Facing the Media. In: Lindstone, J. (ed.) *Presentation Planning and Media Relations for the Pharmaceutical Industry*, pp. 89–91. Routledge, London (2018)
18. Lidstone, J.: Dealing with the press. In: Lindstone, J. (ed.) *Presentation Planning and Media Relations for the Pharmaceutical Industry*, pp. 92–105. Routledge, London (2018)
19. Shannon, C.E., Weaver, W.: *The Mathematical Theory of Communication*. The University of Illinois Press, Illinois (1964)
20. Fiske, J.: *Introduction to Communication Studies*. Routledge, London (2010)
21. Sperber, D., Wilson, D.: *Relevance: Communication and Cognition*. Blackwell, London (1995)
22. Dimpleby, R., Burton, G.: *More than Words: An Introduction to Communication*, 4th edn. Routledge, London (2007)
23. Breton, P., Proulx, S.: *L’explosion de la communication: Introduction des théories et aux pratiques de la communication*. La Découverte, Paris (2016)
24. Monteiro, A.C., Caetano, J.A.C., Marques, C.H., Lourenço, J.M.L.: *Fundamentos de comunicação*. Edições Sílabo, Lisboa (2008)
25. Littlejohn, S.W., Foss, K.A.: *Theories of Human Communication*. Thomson Learning, Boston (2008)
26. Mueller, E.A.: A utilização das ferramentas de comunicação digital na divulgação da atividade profissional de relações públicas pelo coletivo “todo mundo precisa de um RP”. In: Gonçalves, G., Athaydes, A. (eds.) *Relações públicas e comunicação organizacional- Dos fundamentos às práticas: liderança, inovação e empreendedorismo*, pp. 13–30. LabCom.IFP, Covilhã (2017)

27. Morgado, S. M. A., Ferraz, R.: Social media: Facebook e a gestão policial no planeamento de grandes eventos. *Int. J. Market. Commun. New Media* **1**(Special number-QRMCNM), 27–47 (2016)
28. Boyd, D.M., Ellison, N.B.: Social network sites: definition, history, and scholarship. *J. Comput.-Mediat. Commun.* **13**(1), 210–230 (2008)
29. Wohlgenuth, V., Parmeggiani, B.: As organizações e a interação com os usuários no Facebook: o caso da Netflix no Brasil e em Portugal. In: Gonçalves, G., Athaydes, A. (Eds.) *Relações públicas e comunicação organizacional: Dos fundamentos às práticas: liderança, inovação e empreendedorismo*, pp. 31–51. LabCom.IFP, Covilhã (2017)
30. Dias, R., Glerian, A.G., Silva, L.M., Gomes, V.M.L.R.: Como a teoria das representações sociais pode contribuir para o debate sobre “públicos” no campo das relações públicas: Apontamentos exploratórios. *Revista de Humanidades* **32**(2), 250–261 (2017)
31. Novaes, A., de Lourdes Ornellas, M., Ens, R.T.: Convergências teóricas em representações sociais e seu aporte para o estudo de políticas docentes. *Revista Diálogo Educacional* **17**(53), 999–1015 (2017)
32. Zémor, P.: *La communication publique*. PUF (1995)
33. Hatch, M.J., Schultz, M.: The dynamics of organizational identity. *Human Relations* **55**(8), 989–1018 (2002)
34. Gioia, D.A., Schultz, M., Corley, K.G.: Organizational identity, image, and adaptive instability. *Acad. Manag. Rev.* **25**(1), 63–81 (2000)
35. Dutton, J.E., Dukerich, J.M., Harquail, C.V.: Organizational images and member identification. *Adm. Sci. Q.* **39**(2), 239–263 (1994)
36. Barichello, E.M., Dall’Agnese, C.T.: Estratégias de visibilidade e legitimação institucional na produção jornalística hipermédia. *Revista Comunicação Midiática* **11**(3), 105–117 (2016)
37. Smith, R.D.: *Strategic Planning for Public Relations*. Routledge, London (2017)
38. Falkheimer, J., Heide, M.: From public relations to strategic communication in Sweden: the emergence of a transboundary field of knowledge. *Nordicom Rev.* **35**(2), 123–138 (2014)
39. Hopmann, D.N., Skovsgaard, M., Elmelund-Præstekær, C.: What have I done to deserve this? The role of deservingness in effects of ordinary citizens as cases in the news. *Eur. J. Commun.* **32**(4), 333–347 (2017)
40. Bennett, W.L., Livingston, S.: The disinformation order: disruptive communication and the decline of democratic institutions. *Eur. J. Commun.* **66**(2), 122–139 (2018)
41. Steyn, B., Puth, G.: *Corporate Communication Strategy*. Heinemann Publishers, Portsmouth, New Hampshire (2000)
42. PSP: *Grandes Opções Estratégicas da PSP para 2017–2020*. Direção Nacional da PSP, Lisboa (2016)
43. Lindlof, T.R., Taylor, B.C.: *Qualitative Communication Research Methods*. Sage, Thousand Oaks, CA (2011)
44. Daba-Buzoianu, C., Bîră, M., Tudorie, G., Duduciuc, A.: Introductory remarks on inquiring communication through qualitative research. In: Bîră, M., Daba-Buzoianu, C., Duduciuc, A., Tudorie, G. (eds.) *Exploring Communication Through Qualitative Research*, pp. 1–8. Cambridge Scholars Publishing, Cambridge (2017)
45. Morgado, S.M.A., Moniz, T., Felgueiras, S.: Facebook and Polícia de Segurança Pública: An exploratory study if follower’s engagement. In: Rocha, Á., Reis, J., Peter, M., Bogdanović, Z. (eds.) *Marketing and Smart Technologies. Smart Innovation, Systems and Technologies*, pp. 363–376. Springer, Cham (2020)
46. Morgado, S.M.A., Anjos, O.: Qualitative methodology helping police sciences: building a model for prevention of road fatalities in São Tomé and Príncipe. In: Costa, A. Reis, L., Moreira, A. (eds.) *Computer Supported Qualitative Research. WCQR 2018. Advances in Intelligent Systems and Computing*, pp. 292–304. Springer, Cham (2019)
47. Novak, J.D., Cañas, A.J.: The theory underlying concept maps and how to construct and use them. In: *Technical Report IHMC CmapTools 2006–01, Rev 01–2008* (pp. 1–36). Institute for Human and Machine Cognition, Flórida (2008)

48. Daley, B.J.: Concept maps: practice applications in adult education and human resource development. *New Horizons Adult Educ. Human Resour. Dev.* **24**(2–4), 30–36 (2010)
49. Myrdal, G.: *Objectivity in Social Research*. Wesleyan University Press, Middletown (1983)
50. Flick, U.: *The SAGE Handbook of Qualitative Research*. Sage, Thousand Oaks, CA (2017)
51. Quivy, R., Marquet, J., Campenhoudt, L.V.: *Manuel de recherches en sciences sociales*. Dunod, Malakoff (2017)
52. Benn, S., Abratt, R., O’Leary, K.: Defining and identifying stakeholders: views from management and stakeholders. *South Afr. J. Bus. Manage.* **47**(2), 1–11 (2016)
53. Boros, S.: Identity and image: the soul and face of organizations. In: Boros, S. (ed.) *Exploring Organizational Dynamics*, pp. 644–654. Sage Publications, Thousand Oaks, CA (2009)
54. Brown, T.J., Dacin, P.A., Pratt, M.G., Whetten, D.A.: Identity, intended image, construed image, and reputation: an interdisciplinary framework and suggested terminology. *J. Acad. Mark. Sci.* **34**(2), 99–106 (2006)
55. Cunha, N., Ghisleni, T., Muller, F.D.: O: As estratégias comunicacionais da marca Melissa. *Saber Humano* **7**(10), 144–154 (2017)
56. França, F.: Relações públicas: Atividade estratégica de relacionamento e comunicação das organizações com as partes interessadas. *Anuário Unesco/Metodista de Comunicação Regional* **12**(12), 55–70 (2008)
57. Kent, M. L., Taylor, M.: Social media in public relations: reflections on extending and narrowing relationships. In: *The Proceedings of 7th International Forum on Public Relations and Advertising*, pp. 11–18. Mahidol University, Tailândia (2014)
58. Kent, M.L., Taylor, M.: Putting the social back in social media: a longitudinal, meta-analysis of social media research. *Int. J. Interdiscip. Res.* **5**(1), 62–75 (2016)
59. Moreno, J.: O valor económico da informação na sociedade em rede. *Observatório (OBS*) J.* **9**(2), 1–28 (2015)
60. Kotler, P., Keller, K.L.: *Marketing Management*. Pearson Education, London (2012)
61. Galpin, T.J.: *The Human Side of Change: A Practical Guide to Organization Redesign*. Wiley, New Jersey (1996)
62. Morgado, S., Alves, R.: Core capabilities: body-worn cameras in Portugal. *Eur. Law Enforcement Res. Bull.* **18**(Winter 2019), 107–122 (2019)
63. Wright, J.T., Giovinazzo, R.A.: Delphi: Uma ferramenta de apoio ao planeamento prospectivo. *Caderno de Pesquisas em Administração* **1**(12), 54–65 (2000)
64. Thomson, W.A., Ponder, L.D.: Use of Delphi methodology to generate a survey instrument to identify priorities for state allied health. *Allied Health Behav. Sci.* **2**(4), 383–399 (1979)
65. Linstone, H.A., Turoff, M.: *The Delphi method: Techniques and applications*. Addison-Wesley Educational Publishers Inc., Boston (2002)
66. Dimitrijević, B., Simić, V., Radonjić, V., Kostić-Ljubisavljević, A.: The Delphi method as a research tool: an application in transportation and logistics systems evaluations. In: Arsovski, S., Lazić, M., Stefanović, M. (eds.) *Conference Manual 6th International Quality Conference*, pp. 401–406. University of Kragujevac (2012)
67. Dalkey, N.C., Helmer, O.: An experimental application of the Delphi method to the use of experts. *Manage. Sci.* **9**(3), 458–467 (1963)
68. Grisham, T.: The Delphi technique: a method for testing complex and multifaceted topics. *Int. J. Manag. Proj. Bus.* **2**(1), 112–130 (2009)
69. Hasson, F., Keeney, S., McKenna, H.: Research guidelines for the Delphi survey technique. *J. Adv. Nursing* **32**(4), 1008–1015 (2000)
70. Delphi technique: making sense of consensus. *Practical Assess. Res. Eval.* **12**(10), 1–8 (2007)
71. Okoli, C., Pawlowski, S.D.: The Delphi method as a research tool: an example, design considerations and applications. *Inf. Manage.* **42**(1), 15–29 (2004)
72. Fink, A., Kosecoff, J., Chassin, M., Brook, R.H.: Consensus methods: characteristics and guidelines for use. *Am. J. Public Health* **7**(9), 979–983 (1984)
73. Bennett, W.L., Pfetsch, B.: Rethinking political communication in a time of disrupted public spheres. *J. Commun.* **68**(2), 243–253 (2018)

74. Freedman, D.: Populism and media policy failure. *Eur. J. Commun.* **33**(6), 604–618 (2018). <https://doi.org/10.1177/0267323118790156>
75. Bonfadelli, H.: The internet and knowledge gaps: a theoretical and empirical investigation. *Eur. J. Commun.* **17**(1), 65–84 (2002)
76. Newman, A.: Communication planning: a template for organizational change. *Cornell Hospitality Report* **16**(3), 3–6 (2016)
77. Dahlgren, P.: The Internet, public spheres, and political communication: dispersion and deliberation. *Polit. Commun.* **22**(2), 147–162 (2005)

Database Encryption Performance Impact on PostgreSQL and MongoDB



Marcelo Costa, Margarida Rodrigues, Pedro Baptista, Cristina Wanzeller,
Pedro Martins, and Maryam Abbasi

Abstract The security of data is an essential aspect for organizations because today's most significant trade-off is information. Although assuring information is safe is the most critical aspect for organizations, there are other concerns about making information secure. Encryption is very used to protect data, but what are the costs in terms of performance? According to Gomes et al. (Database encryption for balance between performance and security, 2021 [1]), encryption has a significant impact on database read and performance. The impact of encryption in different databases was tested using PostgreSQL and MongoDB. We wanted to understand if it was more beneficial to use encryption at the client side (PostgreSQL) or Rest (MongoDB). The results have shown that using encryption at Rest improved the performance of the inserts in the database.

Keywords Database · Encryption · Performance · PostgreSQL · MongoDB

M. Costa · M. Rodrigues · P. Baptista
Polytechnic of Viseu, Viseu, Portugal
e-mail: estgv16800@alunos.estgv.ipv.pt

M. Rodrigues
e-mail: estgv16061@alunos.estgv.ipv.pt

P. Baptista
e-mail: estgv17058@alunos.estgv.ipv.pt

C. Wanzeller · P. Martins (✉)
CISeD—Research Centre in Digital Services, Polytechnic of Viseu, Viseu, Portugal
e-mail: pedromom@estgv.ipv.pt

C. Wanzeller
e-mail: cwanzeller@estgv.ipv.pt

M. Abbasi
CISUC—Centre for Informatics and Systems, University of Coimbra, Coimbra, Portugal
e-mail: maryam@dei.uc.pt

1 Introduction

Companies heavily rely on information systems, and considering knowledge as an organizational asset, it must be protected.

This work is based on the future work proposed in the “Database Encryption Performance Impact” article, published last year [1]. The main objective in the referenced paper was to prove that encryption had a significant impact on the database read and write performance; this led to the conclusion that it was inevitable to decrease performance in data encryption. As future work, it proposed an analysis about the encryption mode that was tested but with larger volumes of data to continue the evaluation of the performance degradation with the increase of data quantity and the impact of encryption in different databases.

We also decided to explore another possibility. To analyze if performance improved when using encryption at Rest, we used MongoDB. We encrypted the data at Rest—simultaneously, using client-side encryption in PostgreSQL to see the differences between the two types of encryption. The results have shown that using encryption at Rest improved the performance of the inserts in the database.

The remainder of this work is organized as follows. Section 2 presents the state-of-the-art review. Section 3 outlines our methodology. In Sect. 4, we present the final product preview.

Finally, in Sect. 5, conclusions are drawn, followed by the introduction of future work guidelines.

2 Related Work(s)

In this section, we will present some methodologies adopted in similar research. This research is a part of the future work presented in [1]. This paper presented results about a performance analysis for encryption in a relational database, using PGP and AES. The results were performance degradation when using encryption to read data. When using keys, the time to process was less and more space occupied when using encryption or keys. This means it is challenging to avoid loss of performance when using data encryption [1].

The article proposed by M. Zeghid, M. Machhout, L. Khriji, A. Baganne, and R. Turki gives a modified AES algorithm for the encryption of text and images. The authors of this research started with an analysis of the AES algorithm, and then, they added a keystream generator to guarantee a better encryption performance. They concluded that this new scheme could offer higher security, and it is easier to apply it in software and hardware [2].

Another paper we explored was proposed by Shaza D. Rihan, Ahmed Khalid, and Safe Eldin F. Osman. The starting point was the role that encryption algorithms play to secure data. Their main objective was to compare the performance of AES and DES algorithms, considering processing time, CPU usage, and encryption throughput

on Windows and MAC platforms for different text sizes. The results demonstrated that AES is faster than DES in execution time, and it has higher throughput. DES algorithm proved to be better than AES in CPU usage consumption [3].

Eman Salim Ibrahim Harba researched a method to protect data transferring using three hybrid techniques. Symmetric AES was used to encrypt files, asymmetric RSA to encrypt AES, and HMAC was used to encrypt the symmetric password. This combination allowed us to explore the advantages of each one, to build a high-security system. As a result, it had simple and fast encryption while using low computational requirements and providing high system security [4].

The paper's [5] primary purpose was to illustrate how a problem solved using MySQL would behave when MongoDB was being used on a big data dataset. The results are that MongoDB has better performance than MySQL when inserting, updating, deleting, and selecting data from the database.

3 Experimental Setup

This section overviews the analysis of different encryption types and the proposed work to test the impact of encryption in different databases.

3.1 *Encryption at Rest*

In order to understand the concept of encryption at Rest, it is important to distinguish data in motion and data at Rest. Data in motion is considered operational data because it is manipulated daily. Data at Rest is stored and most likely protected by a firewall, for example [6].

One of the most severe problems of MongoDB is that data files are not encrypted by default during transmission [7]. MongoDB Enterprise 3.2 introduces a native encryption option. This feature allows MongoDB to encrypt data files so that only the parts with the decryption key can decode and read the data. The process of encrypting data includes generating a master key and keys for each database, the encryption of data with the database keys, and database keys with the master key [8].

3.2 *AES Encryption Algorithm*

The most popular and used symmetric encryption algorithm likely to be encountered nowadays is the Advanced Encryption Standard (AES). With AES, we can use three different block ciphers. AES-128 uses a 128-bit key length to encrypt and decrypt a block of messages, AES-192 uses a 192-bit key length, and AES-256 has a 256-bit key length to encrypt and decrypt messages. Each cipher encrypts and decrypts data

in blocks of 128 bits using cryptography keys of 128, 192, and 256 bits, respectively [9]. The process of encryption begins with byte substitution, and then, three more steps follow this one, shift rows, mix columns, and add round key, respectively. The process of decryption is similar to the encryption process but in reverse order. The four processes mentioned before are conducted reverse: add round key, mix columns, shift rows, and byte substitution [10].

3.3 Proposed Work

In Fig. 1, we show the solution architecture. For this experiment, 10 GB, 20 GB, and 40 GB of data were generated with the TPC-H. The interaction with the PostgreSQL database was established through Python using the SQLAlchemy module and Pandas to open and read the dataset files. Since the AES algorithm proved to be more efficient for insertion, according to [1], we decided to use it with PostgreSQL.

The interaction with the MongoDB database was also using Python to get the data file. The default encryption mode that MongoDB Enterprise uses is the AES256-CBC (or 256-bit Advanced Encryption Standard in cipher block chaining mode) via OpenSSL. AES-256 uses the same key to encrypt and decrypt text [8]. This means the encryption is not made on the client side, like in PostgreSQL, it will allow us to understand if it is more profitable to perform the encryption at the client side or Rest.

We uploaded the database to a virtual machine in AWS to test the performance with the following features (eight vCPUs, 2.3 GHz, -, 32 GiB memory). Then, we

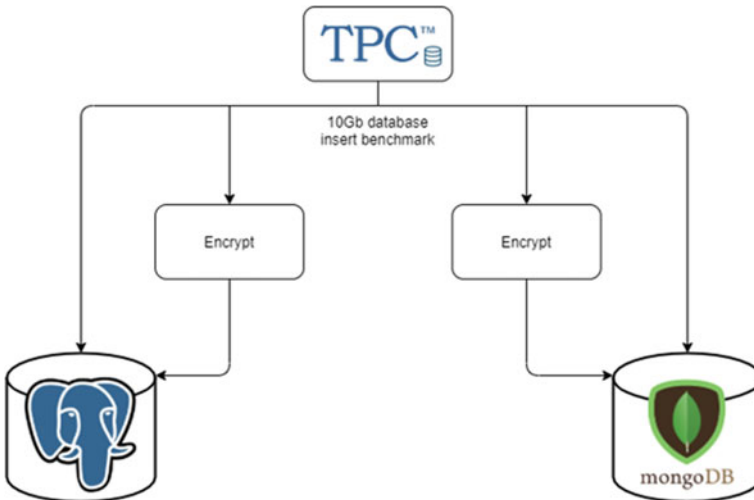


Fig. 1 Solution architecture



Fig. 2 Performance comparison between databases, PostgreSQL and MongoDB, load 10 GB

did the inserts corresponding to 10 GB of data to compare with the database with better performance with and without encryption.

4 Results and Analysis

The test database size was 10 GB in AWS. The results show that MongoDB has a better performance than PostgreSQL when inserting 10 GB of data in the database. Figure 2 shows similar results when it comes to encrypted data, but PostgreSQL has degradation of performance when compared to MongoDB. This means that when encryption is performed on the client side, it will take more time to do the inserts, so MongoDB proved to be a better choice to do the inserts because the encryption is done at Rest.

Figure 3 and Table 1 show the comparison between the inserts of the columns of the database. We did this graphic and table to see if there were significant differences in the times. In some cases, PostgreSQL is slightly quicker than MongoDB, but in the total time spent, we saw that MongoDB had a better performance on the previous analysis. The difference is sharper in tables with more columns, like the table “Line item.”

5 Conclusions and Future Work

When choosing an encryption algorithm suitable for a database, there should be some criteria in consideration. Our work has shown that, although there is a rupture in performance when using encryption, it is possible to overcome this issue by choosing

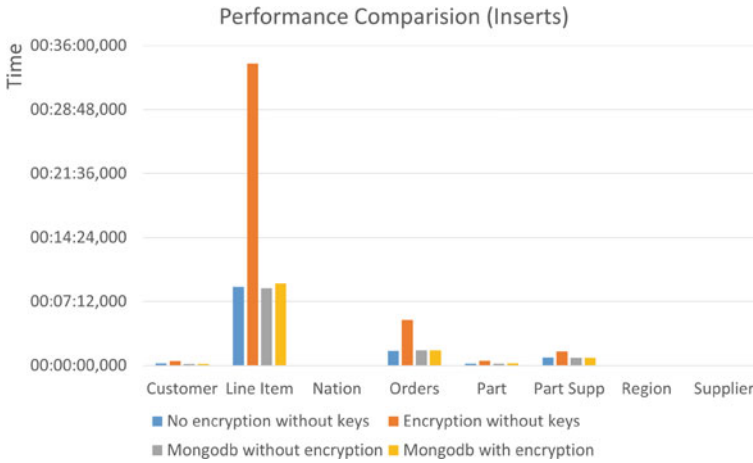


Fig. 3 Performance comparison between databases, PostgreSQL and MongoDB, inserts

Table 1 Performance comparison between databases, PostgreSQL and MongoDB, inserts

	PostgreSQL		MongoDB	
	No encryption	With encryption	No encryption	With encryption
Costumer	00:23:13	00:51:26	00:16:11	00:16:09
Line item	14:45:13	56:35:02	14:30:33	15:23:22
Nation	00:00:00	00:00:00	00:00:00	00:00:01
Orders	02:44:49	08:31:03	02:50:17	02:52:02
Part	00:21:17	00:54:22	00:21:15	00:23:34
Part Supp.	01:28:50	02:40:40	01:26:41	01:26:35
Region	00:00:00	00:00:00	00:00:00	00:00:00
Supplier	00:00:50	00:03:43	00:01:41	00:00:19
Total	19:44:12	69:36:16	19:26:38	20:22:02

other methods and algorithms. Results show that MongoDB has a better performance than PostgreSQL when inserting 10 GB of data in the database. Despite similar results, when it comes to encrypted data, PostgreSQL has degradation of performance when compared to MongoDB. So, when encryption is performed on the client side, it will take more time to do the inserts. MongoDB proved to be a better choice to do the inserts because the encryption is done at Rest.

5.1 Future Work

Although we presented some conclusions about the performance of database encryption when using different databases, this study is not over. We recommend a study with MongoDB where the tests should be performed on the client side to compare with the same database if the best option is to do the encryption at Rest. The sizes 20 GB and 40 GB of data were not tested because we ran out of credit in AWS, so another suggestion would be to test the inserts with this quantity of data. We also recommend some improvements:

- In PostgreSQL, our suggestions are reduced indexes, reduce unique keys, use separate disks for write-ahead logging (WAL) and data, and use parallel writes;
- In MongoDB, reduced indexes, concurrent inserts, and more performing hardware are our recommendations.

Acknowledgements This work is funded by national funds through the FCT—Foundation for Science and Technology, IP, within the scope of the project Ref. UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD), the Polytechnic of Viseu for their support.

References

1. Gomes, A., Santos, C., Wanzeller, C., Martins, P.: Database encryption for balance between performance and security (2021)
2. Zeghid, M., Machhout, M., Khriji, L., Baganne, A., Tourki, R., et al.: A modified AES based algorithm for image encryption. *Int. J. Comput. Sci. Eng.* **1**(1), 70–75 (2007)
3. Rihan, S.D., Khalid, A., Osman, S.E.F.: A performance comparison of encryption algorithms AES and DES. *Int. J. Eng. Res. Technol. (IJERT)* **4**(12), 151–154 (2015)
4. Harba, E.S.I.: Secure data encryption through a combination of AES, RSA and HMAC. *Eng. Technol. Appl. Sci. Res.* **7**(4), 1781–1785 (2017)
5. Chickerur, S., Goudar, A., Kinnerkar, A.: Comparison of relational database with document-oriented database (MongoDB) for big data applications. In: 2015 8th International Conference on Advanced Software Engineering & Its Applications (ASEA), pp. 41–47. IEEE (2015)
6. What is encryption at rest, and why is it important for your business? (2020)
7. Abimbola, J.O., Festus, O.A.: Comparative security vulnerability analysis of NOSQL and SQL database using MongoDB and MariaDB
8. Encryption at rest
9. What is encryption and how does it work? (2020)
10. Advanced encryption standard

Multi-DB Data Streaming on Polyglot Systems



Luis Oliveira, João Brito, Frederico Cá, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi

Abstract Several systems today produce enormous textual, numerical, geospatial, structured, and unstructured data. This data may serve many business requirements, which implies specific performance requirements. These requirements typically make it imperative that the application uses the most efficient means to store and retrieve it. In the past, systems architects would implement SQL relational databases systems (RDBMS). Still, today, the advances of NoSQL and document storage technologies offer a high-scalable non-relational database that can process and store vast amounts of that with high performance and efficiency. Mixing the traditional relational databases with the new databases creates a database system model capable of serving the different business needs. This paper proposes to create a stream architecture capable of feeding the same data to other database systems based on open-source technologies such as Apache Kafka, NoSQL database, time series database (TSDB), relational database, and document indexing engine. The implementation will handle massive incoming data from processed network traffic traces that will be ingested to several databases through the stream architecture.

Keywords Stream · RDBMS · NoSQL · TSDB · Big data · Polyglot persistence

L. Oliveira · J. Brito · F. Cá
Polytechnic of Viseu, Viseu, Portugal
e-mail: estgv14327@alunos.estgv.ipv.pt

J. Brito
e-mail: estgv18530@alunos.estgv.ipv.pt

F. Cá
e-mail: estgv18529@alunos.estgv.ipv.pt

C. Wanzeller · P. Martins (✉)
CISeD - Research Centre in Digital Services, Polytechnic of Viseu, Viseu, Portugal
e-mail: pedromom@estgv.ipv.pt

C. Wanzeller
e-mail: cwanzeller@estgv.ipv.pt

M. Abbasi
CISUC - Centre for Informatics and Systems, University of Coimbra, Coimbra, Portugal
e-mail: maryam@dei.uc.pt

1 Introduction

Massive volumes of textual, numerical, geospatial, structured, and unstructured data are being generated at high rates. Traditional databases such as Oracle, MySQL, PostgreSQL, and other relational databases (RDBMS) were not designed to complete today's data needs. One good example of the unfitted design model of RDBMS is the big data. "Big data is a term to describe a bunch of data with six' v characteristics including volume, velocity, variety, veracity, variability, and value. Big data computing is a popular term these days to solve complex data, and many companies create their extensive data system to handle their massive data" [1]. This big data, or massive data volumes, not address and exposed in a form that can be easily used, may lead to an inability to deliver tangible benefits for the organization or business.

New infrastructural challenges rise to achieve the best efficiency and availability of data through the implementation of architectures based on data streaming and message queuing. "Streaming data refers to data that is continuously generated, usually in high volumes and at high velocity." The message queuing is a fundamental concept to provide the ability to the architecture to store messages and allow asynchronous interaction.

While traditional data solutions are focused on reading and writing data in batches, streaming architectures consume data as it is being generated and may perform data manipulation. Databases, document storage search engines, and other data consumers, visualizers, or alerting systems ingest in real-time upward the consuming pipeline.

The streaming architecture is composed of a message broker, producers, and consumers. The message broker receives data from the producers, transforms it into a standard message format, and streams it. The other components, designated consumers, listen and consume the messages from the message broker.

The first generation of message brokers, such as RabbitMQ and Apache ActiveMQ, relied on the message-oriented middleware (MOM) paradigm, which "provides a clean method of communication between disparate software entities. MOM is one of the cornerstone foundations that distributed enterprise systems are built upon. MOM can be defined as any middleware infrastructure that provides messaging capabilities" [2]. More recently, other high-performance messaging platforms, more suitable for streaming, have emerged. Between then, there are two popular ones, Apache Kafka and Amazon Kinesis Data Streams. "Apache Kafka is a community-distributed event streaming platform capable of handling trillions of events a day. Initially conceived as a messaging queue, Kafka is based on an abstraction of a distributed commit log. Since being created and open-sourced by LinkedIn in 2011, Kafka has quickly evolved from messaging queue to a full-fledged event streaming platform" [3]. Kafka through APIs and connectors allows integration with several applications, databases, and storage systems. Figure 1 illustrates the Kafka environment. "Amazon Kinesis Data Streams (KDS) is a massively scalable and durable real-time data streaming service. KDS can continuously capture gigabytes

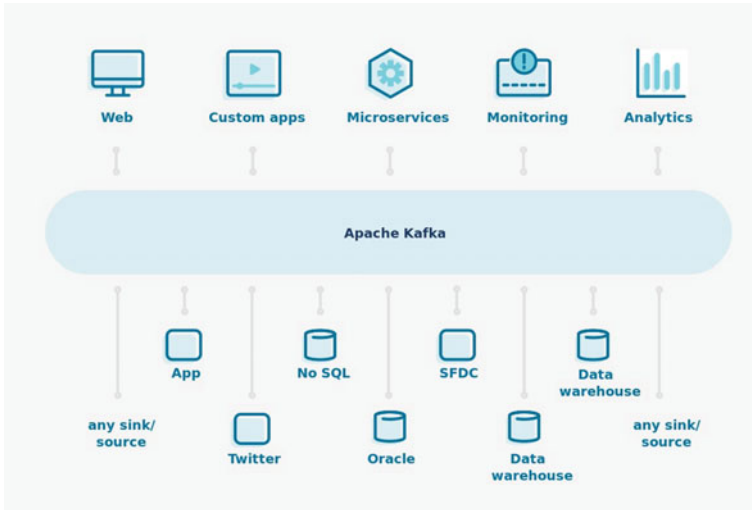


Fig. 1 Kafka interaction [3]

of data per second from hundreds of thousands of sources such as Web site click-streams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events.” [4]

Big data is strongly connected to the concept of distributed systems, in which the CAP theorem (consistency, availability, and partition tolerant) points out the prominent use of the eventual consistency property in distributed systems. This has pushed the need to use different types of databases beyond Structured Query Language (SQL) that have properties of scalability and availability. Not-only SQL (NoSQL) databases, mainly with the basically available, soft state, and eventual consistency (BASE), are gaining ground in the significant data era. In contrast, SQL databases are left trying to keep up with this paradigm shift [5]. Each one of these database types has its strengths and weaknesses, which makes them not perfect. One way to solve this problem is to use a polyglot persistence model. “Polyglot persistence is an enterprise storage term used to describe choosing different data storage/data stores technologies to support the various data types and their storage needs. Polyglot persistence is essentially the idea that an application can use more than one core database (DB)/storage technology” [6]. The streaming architecture plays an essential role in the polyglot persistence’s design and implementation.

The work presented in this paper aims to demonstrate the implementation of a multilingual persistence model based on a producer, a message broker, and several consumers.

The producer used network traffic traces, either online or offline, captured with Wireshark. This network protocol analyzer allows to capture packet data from a live network or read packets from a previously saved capture file. Packages are materialized into a decoded form of those packets to the standard output or writing

the packages to a file [7]. The trace was then piped into a Python script in which data was enriched with geolocation information provided by MaxMind [8] and data ingested to Kafka message broker.

On the consumer side, different databases types were used:

- NoSQL:
 - MongoDB, which is a general purpose, document-based, distributed database built for modern application developers and for the cloud era [9].
 - InfluxDB, a time series database (TSDB) [10].
- Document-oriented database: ElasticSearch, which is the world's leading free and open search and analytics solution [11].
- Relational database: PostgreSQL, which is an open-source object-relational database [12].

Besides the previously described implementation, a multi-platform open-source analytic and interactive visualization web application was also used, Grafana. It provides charts, graphs, and alerts when connected to supported data sources [13].

The complete experimental work setup was performed on a single server running Docker [14] with multiple containers installed with images obtained from the Docker Hub [15].

With this experimental work, we were able to demonstrate a fully functional multilingual persistence model system. The data goes through a single application, distributed to several consumers from a single point (broker), keeping the consistency between all the different databases and at the same time delivering a system capable to respond to several business needs.

The paper is organized as follows. Section 2 presents the state-of-the-art review. Sect. 3 presented the architecture and technologies used in the experimental work. Section 4 describes the experiment setup. The experiment work and results analysis are presented in Sect. 5. Finally, in Sect. 6, conclusions are drawn, followed by the introduction of future work guidelines.

2 Related Work(s)

The increasing needs for scalability and performance have gradually pushed new alternative database systems, namely NoSQL technology, gradually reducing traditional relational databases. With this change, enormous studies and papers have been written. Some work focuses on the performance comparison between SQL and NoSQL databases operations of reading, write, delete, and instantiate processes on key-value stores [16]. Others concentrate on choosing the best NoSQL database for a specific job but taking into account quality attribute focused survey or classification of NoSQL databases where databases are compared with regards to their suitability for quality attributes common on the design of enterprise systems [17]. More works

were done on performance, based on CAP’s theorem, of streaming data processing using popular NoSQL databases [18].

3 Architecture and Technologies

Aiming to demonstrate a multilingual system and evaluate results, a fully functional experiment involving a message broker and data streamer able to feed several databases types through it was implemented. The architecture and technologies presented in Fig. 2 involve concepts of producers, message brokers, and consumers as central players in this implementation. The producer is composed of a Python script receiving piped data from a network protocol analyzer designated Wireshark [7]. To enrich data piped by the Wireshark, a GeoIP database provided by MaxMind [19] in conjunction with a Python API [20] was used. The message broker and data streamer are composed of Kafka by Confluent [3]. Kafka stands as the orchestrator of the complete architecture by interacting with the producer and consumers.

Composing that the consumers are tree different databases types and a document indexing search engine:

- MongoDB and influxdb can be categorized as NoSQL databases. The first is a general purpose, document-based, distributed database [9]. The second one is a time series database (TSDB) [10].
- PostgreSQL, which is an open-source object-relational database [12].
- ElasticSearch, document indexing search engine is the fourth consumer element of the architecture [11].

Finally, with the purpose of easily visualize the data stored in the different databases, additional applications were used:

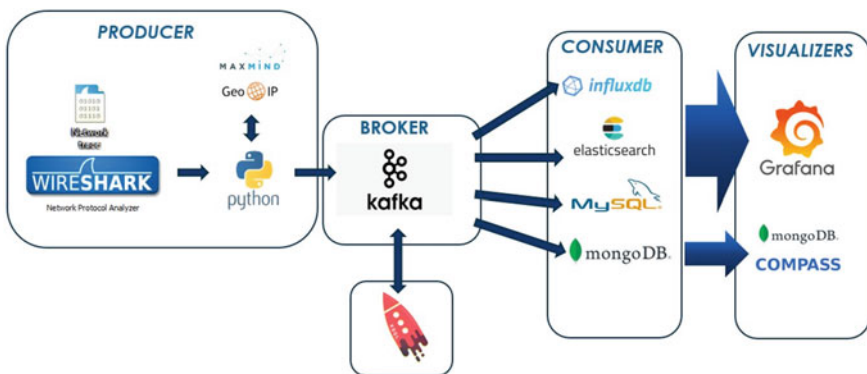


Fig. 2 Experiment architecture

- Grafana, which is multi-platform open-source analytics and interactive visualization web application that provides charts, graphs, and alerts when connected to supported data sources [13].
- MongoDB COMPASS, a GUI for MongoDB, allows visualizing and exploring data on the database. Additionally, it also allows to run ad hoc queries in seconds and interact with the data with full CRUD functionality [21].

4 Experimental Setup

The experiment was run in phases, and each step brought more complexity to the investigation.

The first phase started by implementing the base system. This base system was composed of the producer. As previously mentioned, it contained a network protocol analyzer, a Python script, and a GeoIP database, followed by the implementation of the message broker streamer. For faster deployment, this and the following phases were done using a containerized solution using Docker images. In this phase, a set of tests was performed using different sizes of previously captured network traffic traces files and performing real-time network traffic capture and ingestion.

In the second phase, two consumers were added to the experiment. First, it was added a time series DB (influxDB) and relational DB (MySQL). In this phase and the next phases, new sets of tests were done using a previously captured network trace file containing 509 K packets with an average capture rate of 8.8 packets/s. Table 1 shows the complete trace file properties.

The third phase adds two more consumers to the experiment architecture, Elasticsearch and MongoDB. The first one is the indexing search engine, and the second is a document-oriented database.

In the fourth phase, to test inter-database consistency, a network connectivity interruption to the message broker was simulated.

The final phase was done by integrating ksqlDB, which enables “to build modern, real-time applications with the same ease and familiarity of building traditional applications on a relational database. It also simplifies the underlying architecture for these applications, so you can build powerful, real-time systems with just a few

Table 1 Dataset

Trace file statistics	
Measurement	Captured
File name	dataset.pcap
Packets	509 K
Data byte rate	33 MBps
Data bit rate	271 Mbps
Average packet size	3846.58 bytes
Average packet rate	8,809 packets/s

Table 2 Ingestion combinations

Test no.	Trace file statistics	
	Num. events	Events/s
001	10,000	25
002	50,000	50
003	100,000	100

SQL statements.” The ksqlDB introduced on the experiment near-real-time data enrichment.

Additionally, to the mentioned phase, the data visualizers were also implemented.

5 Experimental Work and Results

5.1 Phase 1: Producer and Message Broker

The main goal of this phase was to verify the performance of the implemented architecture and technologies without any consumer integration. For that, it was defined a set of data ingestion combinations described in Table 2 and used real-time network traffic capture and ingestion. The event data rate output was performed by introducing delays on the Python script.

No data events loss or delays were observed when using any of the ingestion combinations.

5.2 Phase 2: Producer, Message Broker, and Two Consumers

Phase 2 introduces two consumers, influxDB and MySQL. The introduction of the consumers into the Kafka ecosystem was done by using Kafka connectors. These connectors are divided into two types, Source and Sink, and they allow to receive data from other DBs into Kafka or send data out of Kafka into DBs.

Running again the dataset combination scenarios on Table 2, once again, we do not observe any delay or event loss on the events received from the producer. The variable delay was observed on the consumer between the time the events were received by the message broker from the producer and received by the consumers. This delay was observed on the Apache Kafka GUI [22] reported on the column “messages behind,” illustrated on Fig. 3.

After the producer finished sending data to the broker and the consumers, influxDB and MySQL finished to receive all the events. Both databases had the same number of events.

Consumer group ID	Messages behind	Number of consumers
connect-Nettraffic_influxDB	4,180	1
connect-Nettraffic_MySQL	2,222	1

Fig. 3 Two consumers delay

Consumer group ID	Messages behind	Consumer group ID	Messages behind
connect-Nettraffic_influxDB	2	connect-Nettraffic_influxDB	2
connect-Nettraffic_ElasticSearch	106,995	connect-Nettraffic_ElasticSearch	0
connect-Nettraffic_MySQL	0	connect-Nettraffic_MySQL	0
connect-Nettraffic_mongoDB	106,995	connect-Nettraffic_mongoDB	0

(a) Before

(b) After

Fig. 4 Four consumers

5.3 Phase 3: Producer, Message Broker, and Four Consumers

In Phase 3, two more consumers were added, ElasticSearch and MongoDB, which brought higher complexity and experiment resiliency. Using the same dataset ingestion scenarios, no delay or event loss was observed on the message broker related to events coming from the producer. Since there were already events on the topic stored on the Kafka message broker from previous phases, all of these events were sent to the newly introduced consumers. Since these new consumers were subscribing to an existing topic, by the Kafka design, this was expected behavior. This design is based on the principle that any new consumer, or existing, that has not transferred all the events that are inside the retention window will do so until the end offset is equal to the current offset. Figure 4 shows the number of messages behind, not consumed, before and after the two new consumers were activated.

5.4 Phase 4: DB Isolation and Connectivity Restore

The goal of Phase 4 was to test the inter-DB consistency on the polyglot persistence system. This was done by putting the MySQL DB sink connector into pause status and proceeding with the producer's data ingestion. All the other consumers consumed the new just ingested events. Since the MySQL connector was in pause status, simulating

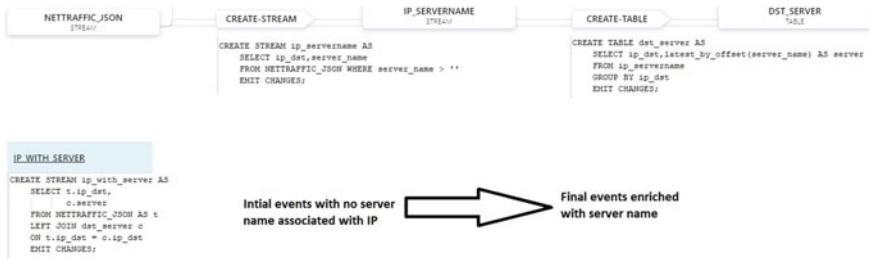


Fig. 5 ksqlDB stream flow

a lack of connectivity or DB problem, MySQL DB was not updated. To “restore connectivity,” the MySQL connector was put in the play status. After some seconds, it was observed that all the missing event records were fed to the DB and the MySQL DB had the same number of events as the other DBs.

5.5 Phase 5: ksqlDB Data Enrichment

Aiming to evaluate the near-real-time data enrichment, this was done using the ksqlDB stream feature of the Confluent Kafka.

From the existing topic, a new data stream was created with only two fields, ip dst and server name. This new stream allowed creating a table containing only the association of the destination IPs and server names. Joining this table with the near-real-time stream provided a near-real-time data enriched stream containing the IP address associated with the server name. Figure 5 illustrates the enriched stream creation.

6 Conclusions and Future Work

The experimental work performed, and described in this paper, allows us to conclude that an architecture containing a message stream broker, a producer capable of ingesting events at a high data rate, and different types of databases, as consumers, has excellent performance and excellent data resilience and persistence. This type of architecture fits quite well in scenarios where the same data has to be store enriched, aggregated, correlated, or in full detail, responding this way to several business needs and retention intervals. Additionally, we may also conclude that the open-source applications available today allow us to construct complex architectures capable of responding to scalability needs in the significant data era. Future work may be conducted to explore new data enrichment and real-time metrics calculations at the

message broker before being fed to the databases. The introduction of more producers with other data types and aggregated into a single stream would also be valuable.

Acknowledgements This work is funded by National Funds through the FCT—Foundation for Science and Technology, IP, within the scope of the project Ref. UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISED), the Polytechnic of Viseu for their support.

References

1. Hamami, F., Dahlan, I.A.: The implementation of stream architecture for handling big data velocity in social media. *J. Phys. Conf. Ser.* **1641**, 012021 (2020)
2. (pdf) message-oriented middleware. https://www.researchgate.net/publication/220035284_Message-Oriented-Middleware. Accessed on 24 May 2021
3. What is apache kafka?—confluent. https://www.confluent.io/what-is-apache-kafka/?utm_medium=sem&utm_source=google&utm_campaign=ch.sem_br.nonbrand_tp.prs_tgt.kafka_mt.xct_rgn.emea_lng_eng_dv.all&utm_term=apache%20kafka&creative=&device=c&placement=&gclid=Cj0KCQjwna2FBhDPARIsACAEC_WjuVPIkETwHVnARFuF0G3cgIqzqyIMcJe1mAntoiGqlcmGVyv-KuQaAvRSEALw_wcB. Accessed on 24 May 2021
4. Amazon kinesis data streams—data streaming service—amazon web services. https://aws.amazon.com/kinesis/data-streams/?nc1=h_ls. Accessed on 25 May 2021
5. Khine, P.P., Wang, Z.: A review of polyglot persistence in the big data world. *Information* **10**(4) (2019)
6. What is polyglot persistence? Definition from whatis.com. <https://searchapparchitecture.techtarget.com/definition/polyglot-persistence>. Accessed on 26 May 2021
7. tshark—the wireshark network analyzer 3.4.5. <https://www.wireshark.org/docs/man-pages/tshark.html>. Accessed on 26 May 2021
8. GeoiP® databases & services: Industry leading IP intelligence—maxmind. <https://www.maxmind.com/en/geoiP2-services-and-databases>. Accessed on 26 May 2021
9. The most popular database for modern apps—mongodb. <https://www.mongodb.com/>. Accessed on 26 May 2021
10. Influxdb time series platform — influxdata. <https://www.influxdata.com/products/influxdb/>. (Accessed on 05/26/2021).
11. Porque gratuito e aberto?—elastic. https://www.elastic.co/pt/about/free-and-open?ultron=fao-sitelink&gambit=Elasticsearch-Core&blade=adwords-s&hulk=cpc&Device=c&thor=elasticsearch&gclid=CjwKCAjw47eFBhA9EiwAy8kzNJ8-4cVDEAnbKrcc6f1EX7baujXEAVuBNacrlom3PlzXxra-DCFF7RoCyUgQAvD_BwE. Accessed on 26 May 2021
12. PostgreSQL: The world's most advanced open source database. <https://www.postgresql.org/>. Accessed on 26 May 2021
13. Grafana—wikipedia. <https://en.wikipedia.org/wiki/Grafana>. Accessed on 26 May 2021
14. Empowering app development for developers—docker. <https://www.docker.com/>. Accessed on 26 May 2021
15. Docker hub. <https://hub.docker.com/>. Accessed on 26 May 2021
16. Li, Y., Manoharan, S.: A performance comparison of SQL and NOSQL databases. In: 2013 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM), pp. 15–19 (2013)
17. Choosing the right nosql database for the job: a quality attribute evaluation — springerlink. <https://link.springer.com/article/10.1186/s40537-015-00250#citeas>. Accessed on 27 May 2021
18. Doguc, T.B., Aydin, A.A.: Cap-based examination of popular NOSQL database technologies in streaming data processing. In: 2019 international artificial intelligence and data processing symposium (IDAP), pp. 1–6 (2019)

19. Geop2 databases—maxmind. <https://www.maxmind.com/en/geop2databases>. Accessed on 29 June 2021
20. Github—maxmind/geop2-python: Python code for geop2 webservice client and database reader. <https://github.com/maxmind/GeoIP2-python>. Accessed on 29 June 2021
21. Compass—mongodb. <https://www.mongodb.com/products/compass>. Accessed on 29 June 2021
22. Apache kafka gui management and monitoring-confluent. <https://www.confluent.io/product/confluent-platform/gui-driven-managementand-monitoring/>. Accessed on 3 July 2021

OSSpal Qualitative and Quantitative Comparison: Couchbase, CouchDB, and MongoDB



Tiago Cardoso, João Penela, André Rosa, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi

Abstract With the rapid increase in cloud computing, big data, and web applications in real time, there is a great need to turn to database platforms that can handle and process large volumes of data, with high performance and excellent scaling capabilities. Since its creation, this has been one of the main goals of NoSQL databases. This article will cover three of the most popular NoSQL databases, MongoDB, Couchbase, and CouchDB. We will take a more theoretical approach to each database and compare them using the OSSpal methodology.

Keywords Big data · NoSQL databases · MongoDB · Couchbase · CouchDB

1 Introduction

There is a large volume of data generated by millions of users who use applications or work platforms daily, whether for personal or professional purposes. With the different requirements of these applications, such as scalability and the high degree of availability, they have contributed to new technologies and platforms. Social

T. Cardoso · J. Penela · A. Rosa
Polytechnic of Viseu, Viseu, Portugal
e-mail: estgv18581@alunos.estgv.ipv.pt

J. Penela
e-mail: estgv18572@alunos.estgv.ipv.pt

A. Rosa
e-mail: estgv19301@alunos.estgv.ipv.pt

C. Wanzeller · P. Martins (✉)
CISeD—Research Centre in Digital Services, Polytechnic of Viseu, Viseu, Portugal
e-mail: pedromom@estgv.ipv.pt

C. Wanzeller
e-mail: cwanzeller@estgv.ipv.pt

M. Abbasi
CISUC—Centre for Informatics and Systems, University of Coimbra, Coimbra, Portugal
e-mail: maryam@dei.uc.pt

networks, for example, require the management of large amounts of unstructured data, which are generated daily by millions of users in search of sharing information, knowledge, and interests. In this context, a new category of database emerged, called Not-only SQL (NoSQL), was proposed to meet the management requirements for large volumes of data, semi-structured or unstructured, which require high availability and scalability. The need for a new BD technology arose due to the inefficiency of relational databases in dealing with this task. NoSQL databases use a flexible data model with no predefined schema. This is the complete opposite of a relational database perspective, with a fixed schema and table structure. This means that each row in a table is defined in a very narrow way in the relational database. Therefore, one of the advantages of NoSQL is that it supports unstructured data, which means that there are no database-level restrictions on what types of data can be stored at any point. Another problem that NoSQL databases can help with is scaling. NoSQL databases are scaled horizontally, as opposed to the relational database, which is scaled vertically. We will describe this in more detail later. Because of the growing interest in adopting this technology and the new challenges generated by the use of NoSQL, it is essential to know the main concepts and their uses. In particular, this knowledge has great relevance for those interested in collaborative Web applications, which, most of the time, need a technology that supports the management and scalability of large volumes of data, simply and effectively. The OSSpal methodology used in this article to compare databases proved to be suitable for collaborative open-source projects. The latter one compares databases in six categories: general quality, instability, usability, robustness, security, and scalability.

2 Related Work

2.1 *NoSQL*

SQL databases known for being relational databases were developed in the 70s to solve data storage and to help with data loss problems. In this model, the data is stored in the rows of a table, if there is a need to add or change a data type (table), we must change the entire database, and in these databases, the scale is vertical and performance can be improved with updating servers and machines [13]. The NoSQL concept was introduced in 1998, and it has an unrelated model that aims to simplify databases while improving performance and flexibility. The scale in NoSQL models is horizontal, if more performance is needed, the system will spread the database to more machines and spread the tasks among them. Due to its simple structure, the performance in the reading and writing processes is very good, combined with high availability, easy scalability and flexibility [14]. Most of these NoSQL systems will fall into one of the following four categories:

- Document-oriented storage: This database model stores your data in the form of a key-indexed document. It offers great performance options and horizontal scalability, which allows them to store more data while maintaining a high performance level [12].
- Key-value storage: A simple yet powerful and efficient model consisting of an application programming interface (API) allows the user to store data in a schema-free form [8].
- Column-oriented database: Somewhat similar to a relational database, this model also stores data in tables, but each row contains a dynamic number of attributes.
- Other non-relational database models, such as a graph database or an object-oriented database, are also considered to be part of the NoSQL category.

2.2 *NoSQL Scaling Capabilities*

As the volume of data increases, the need for scalability and performance improvement increases as well. As mentioned before, one of the characteristic grids of NoSQL systems is horizontal scalability [13]. The absence of locks is critical in NoSQL databases, allowing for horizontal scalability and making this technology suitable for solving exponentially growing data volume management problems. A well-known alternative to achieve horizontal scalability is sharding, which consists of dividing data into multiple tables to be stored across several nodes of a network. Another reason NoSQL models have great scalability is the lack of schema [6]. This lack of schema both facilitates scalability and contributes to increased availability. These structures are, for the most part, based on a key-value concept, allowing for high flexibility in the way data is organized. Another way to promote scalability is through replication. Allowing replication natively cuts down on time spent retrieving information. NoSQL databases use APIs to facilitate access to information, so that any application can use database data quickly and efficiently, to promote an efficient way of accessing data, offering high availability and scalability, that is, the focus is not on how the data is stored, but on how we can retrieve it efficiently [17]. Eventual consistency is also a characteristic of NoSQL databases related to the fact that consistency is not always maintained between the various data distribution points. This characteristic is based on the consistency, availability, and partition tolerance (CAP) theorem, which says that, at a given time, it is only possible to guarantee two of three properties between consistency, availability, and partition tolerance.

2.3 *Document-Oriented Storage*

In this work, we will focus only on tree document storage databases where documents depend on the internal structure to extract data that the database engine uses for optimization. Document stores store the data of an object, that is, an instance of

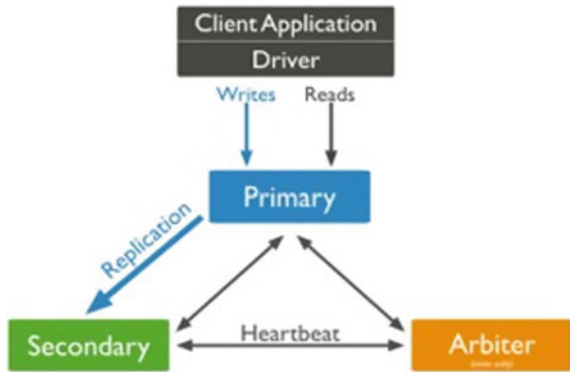
the database, and each object stored can be different from all the others. This eliminates the need for object-relational mapping when loading data into the database [17]. Documents are addressed in the database via a unique key that represents that document. This key is a simple identifier that can be used to retrieve the document from the database. Typically, the database maintains an index on the key to speed up document retrieval, and in some cases, the key is needed to create or insert a document into the database. Document implementations include a variety of document forms, including: collections, tags or metadata, and selection hierarchies. These databases support an API or query language that allows the user to retrieve documents based on content. For example, we might want a query that retrieves all documents with a certain field set to a certain value. To update or edit a document's data, the database allows for document replacement or a structure piece of the document [12, 17].

3 MongoDB, Couchbase, and CouchDB

MongoDB is an open-source NoSQL database management system developed in C++ by MongoDB, Inc. To this day, it has high availability through replication and fault tolerance, horizontal scalability, end-to-end security: to-end, global support, and management tools for automation, monitoring, and backup. MongoDB stands out in the document-oriented segment. The grouping of documents is called the collection and storage format of BinaryJSON (BSON) [14]. By default, MongoDB uses a dynamic schema, which means that documents in a specific collection do not need to have the same set of fields and data types. However, it is possible to implement JSON schema validation. MongoDB also does not implement a query language similar to SQL [9]. MongoDB uses MapReduce to improve database performance [16]. MapReduce is a command that makes available to condense a large amount of data into valid aggregated results. Indexes are unique data structures that condense a small portion of the data in a collection to help improve query performance. Server-side scripts are implemented through JavaScript executions [9]. MongoDB clusters, Fig. 1, are made up of a primary node, secondary nodes, and an arbitrary node [16]. The primary node has the function of receiving all write operations. To provide data redundancy and high availability, MongoDB implements clustered replication. The secondary nodes will have a replica of the data on the primary node and will also apply the same write operations as the primary node. The arbiter node will select a new primary node in case of failure and will not have any data.

The MongoDB cluster has the following components: shard, mongos, and config server. The first, shard, is responsible for storing documents or parts since a collection can be partitioned into multiple shards. The contents of a shard are stored in a set of replicas, a replica set, which consists of a set of hosts that replicates all the information between them. The second, mongos, plays the role of the query router, that is, the interface between the application and the MongoDB cluster. Finally, the third config server is responsible for storing cluster-specific metadata and configurations. Content stored on the config server must also be in a replica set [2]. Replication follows

Fig. 1 MongoDB clusters

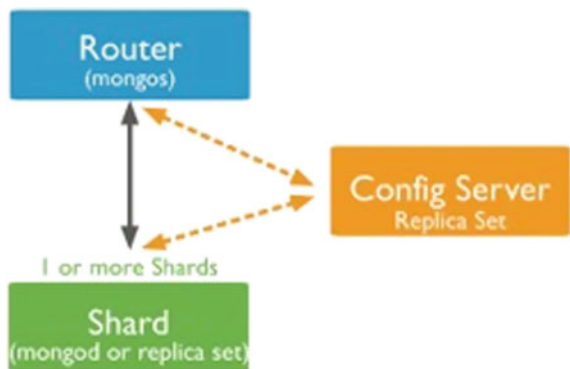


the master–slave mechanism and is identified as a replica set in MongoDB. When configuring a replica set, all nodes that will be part of the cluster shard are informed. The master receives all write requests and replicates them to cluster members asynchronously. In case of problems with the master, automatic failover occurs, in which one of the slaves takes over as master in the replica set [2] (Fig. 2).

CouchDB

CouchDB is a document-oriented open-source database management system that can be accessed through the RESTful JavaScript Object Notation (JSON) API [15] and has drivers for several programming languages. CouchDB was designed to be a new type of database, based on scalability challenges and developed for the web. This database system stores information in JSON documents and does not force the use of a rigid information structure or schema, provides a web interface via HTTP for data access and database management, and allows you to search and index documents using JavaScript. It works with web and mobile applications allowing efficient distribution and replication of data, supporting master-master configurations with automatic conflict detection [5]. CouchDB brings many features, making

Fig. 2 MongoDB sharding configuration



application development easier. It is highly available, fault tolerant, and eventually consistent. It also implements security mechanisms. This database system was used to store information as it is a system that meets the requirements and is easy to install. Another key factor in this decision was the Futon functionality to visualize and manipulate visualizations and data. Couchbase compatibility was also a feature of this system. It was then possible to present an information storage solution suited to the project's requirements. CouchDB uses the slogan "relax" because of its ease of installation. In just a few steps, you can install a document-oriented NoSQL-distributed database system and then development. It is possible to install this system in a cluster environment on several operating systems, from Linux, Mac OSX, Windows, and even mobile devices [6]. CouchDB is a peer-based distributed database system that allows users and servers to access and edit shared data offline and later replicate (bidirectional) those changes across multiple nodes [5].

Couchbase is open-source NoSQL software document-oriented databases. It is based on an architecture that supports flexible JSON templates, easy to scale, high performance, mobile synchronization, and high security. It is the first NoSQL database to enable the development with agility and operate at any scale. Couchbase database emerges from the merge of two products; several leaders of the Memcached project (a distributed memory caching system) founded Membase, and some of these leaders merge with people of CouchDB (distributed memory caching system of apache) in order to create Couchbase server. Couchbase is a memory first solution (distributed cache and the backend are the same), and each time you interact with Couchbase, you write or read directly from memory. From CouchDB, Couchbase has inherited the key-value architecture, and each value has a maximum size of 20 MB for the Couchbase key-value store. However, the most typical way is to use Couchbase like a document store (a document is a JSON file); in this way, you can access the documents with key-value, with index access, or with N1QL (it is an implementation of SQL in JSON). Couchbase has two main products, Couchbase server and Couchbase mobile (an implementation of Couchbase server for embedded systems). Moreover, Couchbase can work as a distributed database as well [4]. Couchbase, with its simple and intuitive interface, provides telemetry from a pre-selected server. Thus, it allows easy visualization of the server's performance and the database, with dynamic and interactive graphics updated with the RAM used in real time, reading, writing, search time, and other helpful information for a database manager. It is also possible to identify other features characteristic of this software, such as buckets, documents, and indexes. Buckets are the elements that connect to applications and store their information in JSON documents [11]. Documents can be compared to records, and they store information in JSON format. Indexes identify the data inside the buckets, and they are a crucial element for running queries efficiently. Couchbase offers a wide variety of indexes and management options. It is also possible to identify shortcuts to management features such as logs, security, server management, and query editor [18].

4 OSSpal Results and Analysis

In this section, we evaluate and compare MongoDB, CouchDB, and Couchbase using the OSSpal methodology to classify the systems. This methodology aims to help organizations find the best quality free software. It is also a widely accepted methodology to combine quantitative and qualitative measures and, consequently, better classify free software [1]. In order to classify the systems, we defined six criteria based on OSSpal: instability, usability, robustness, security, scalability, and overall quality. For each criterion, we defined a weight according to our considerations and research. For the most relevant categories with greater weights, it should be noted that we chose each weight to assign and that in other scenarios, the weights of each category may be different. We gave more weight to global quality, representing thirty percent of the total. We consider this category a global assessment for the system, an aggregation of the other categories, and, therefore, the most influential for the final assessment. Usability with twenty percent of the weight is a critical factor for the success of a system, and for that, we defined a weight greater than instability and security. Robustness and scalability are the categories with less weight, but they are no less important than the previous ones; we are just considering it less relevant for the evaluation of the system (Table 1).

In order to rank each system category, we again followed the OSSpal methodology, assigning a value from zero to five to each category, starting with zero for the lowest score and five for the highest. The evaluation of each system, Table 2, was based on our own experience working with each system the information we researched, ratings, reviews, and analysis in G2 [10], DB-engines [7], and the article [3].

Table 1 Categories weight

Categories	Weights
Instability	15
Usability	20
Robustness	10
Security	15
Scalability	10
Overall quality	30

Table 2 OSSpal results

Categories	MongoDB	Couchbase	CouchDB
Instability	4.5	3.5	3
Usability	5	4.5	4
Robustness	4.5	4	4
Security	4	3.5	3
Scalability	4	4	3.5
Overall quality	4.5	4	3.5

In Table 2, we attribute 4 for robustness in CouchDB and 4 in Couchbase because both of them support live cluster topology changes and 4.5 to MongoDB because it uses BSON files which allow read and write operations faster. All three systems are designed to be as scalable as possible, and both of them are fairly scalable, but by comparison, we have assigned a similar score to all of them with the exception of CouchDB which we give a lesser score, which is 3.5. In terms of instability, we can see that MongoDB stood out a somewhat more, and that is why we gave it a score of 4.5. In terms of usability, we think MongoDB is the best, and because of that, we gave him the maximum score of 5 values, as it has a lot of features and a good user interface. To assess the security and stability of these systems, we resorted to comparisons made in [14] and [15] that once again MongoDB came out to win.

CouchDB: $3 * 0.15 + 4 * 0.2 + 4 * 0.1 + 3 * 0.15 + 3.5 * 0.1 + 3.5 * 0.3 = 3.5$

Couchbase: $3.5 * 0.15 + 4.5 * 0.2 + 4 * 0.1 + 3.5 * 0.15 + 4 * 0.1 + 4 * 0.3 = 3.95$

MongoDB: $4.5 * 0.15 + 5 * 0.2 + 4.5 * 0.1 + 4 * 0.15 + 4 * 0.1 + 4.5 * 0.3 = 4.475$

To compile the information from Table 2 and Fig. 3 and obtain a final value that represents the ranking of each system, we use a mathematical method. The value of each category is multiplied by its weight, and all results are added together to obtain a final system rating. As we can see, MongoDB is the application that obtained the best final score with the application of the OSSpal methodology, with a final score of 4.475 (from 1 to 5), Couchbase gets the score of 3.95, and then, CouchDB gets the worst score of 3.5.

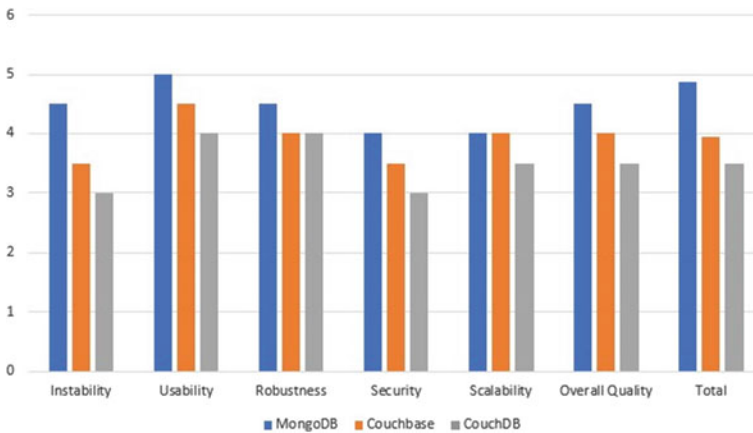


Fig. 3 Graph of final results

5 Conclusions and Future Work

By our analysis, both qualitatively and quantitatively, through the OSSpal methodology, we conclude that MongoDB is the NoSQL database management system (DBMS) open-source with the best results. MongoDB scored 4.475 on our scale, Couchbase and CouchDB 3.95 and 3.5, respectively. CouchDB was the database that we scored the least as it has a big disadvantage compared to Couchbase, and MongoDB is that it uses arbitrary queries, which are expensive in terms of performance, because these queries need a temporary view to run. An advantage we noticed of Couchbase over other systems is that it uses SQL as the query language, which makes it easier for the user to get the data/information they want, and we gave it a score of approximately 4. In general, MongoDB has better scores in almost all categories, because, according to our analysis and uses given, it is the system that offers better resources with higher quality in relation to the others databases. It is essential to mention that based on our results, we cannot say that MongoDB is the best solution for every company. The aim will be to provide more information to help choose the DBMS that best suits each organization's needs. We also have to take into account that these applications were developed for different types of systems that this evaluation is done in windows and that can influence the final result.

As future work, we recommend comparing more NoSQL databases using the OSSpal methodology. It would also be essential and beneficial to test different systems with other benchmarking methodologies to evaluate these applications through their performance in each operation (data creation, update, and deletion), and these tests can take into account the number of records, number of operations per second, and the number of threads.

Acknowledgements This work is funded by National Funds through the FCT—Foundation for Science and Technology, IP, within the scope of the project Ref. UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD), the Polytechnic of Viseu for their support.

References

1. Balaguer, F., Di Cosmo, R., Garrido, A., Kon, F., Robles, G., Zacchiroli, S.: DiCosmo. Open source systems: towards robust practices. In: International Conference OSS 2017, Buenos Aires (2017)
2. D. R. M. J. E. BOVO, Danilo. Avalia, caõ da Replica, caõ e Tolerância a Falhas nos Sistemas Cassandra e MongoDB. WORKSHOP DE TESTES E TOLERANCIAÀ FALHAS (WTF) (2018)
3. Calçada, A., Bernardino, J.: Evaluation of couchbase, couchdb and mongodb using osspal. In: Conference (2019)
4. CouchBase. NoSQLWhite Paper (2021). <http://www.couchbase.com/sites/default/files/uploads/all/whitepapers/NoSQL-Whitepaper.pdf>
5. CouchDB. couchdb.apache.org (2021). <https://couchdb.apache.org>
6. Cutler, J., Dickenson, M.: NoSQL Databases, pp. 193.4.5 (2020)

7. db engines. Couchbase vs. couchdb vs. mongodb comparison (2021). <https://dbengines.com/en>
8. DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Voshall, P., Vogels, W.: Dynamo: Amazon's Highly Available Key-Value Store. *Dynamo: Amazon* (2007)
9. J. e S. Ghemawat. processamento de dados simplificado em grandes clusters. *Comunicações da ACM* (2008)
10. g2. Couchdb, couchbase server, and mongodb. <https://www.g2.com/compare/couchdb-vs-couchbase-couchbase-server-vs-mongodb>. Accessed on 11 June 2021
11. Hubail, M.A., Alsuliman, A., Blow, M., Carey, M., Lychagin, D., Maxon, I., Westmann, T.: Couchbase analytics: Noetl for scalable NoSQL data analysis. In: *Proceedings of the VLDB Endowment* (2019)
12. Sareen, P., Kumar, P.: NoSQL database and its comparison with SQL database. *Int. J. Comput. Sci. Commun. Networks* (2015)
13. Lourenço, J.R., Abramova, V., Vieira, M., Cabral, B., Bernardino, J.: NoSQL databases: a software engineering perspective. *Adv. Intell. Syst. Comput.* (2015)
14. V. C. B. Lourenço, J.R., Abramova. Nosql databases: a software engineering perspective. In: *New Contributions in Information Systems and Technologies* (2015)
15. R. A. e. C. Manyam, Payton. *Relaxe com o couchdb - na era dbms não relacional da bioinformática. Genômica* (2012)
16. Mongo. *Mongo replications doc* (2021). <https://docs.mongodb.com/manual/replication/>
17. Nayak, A., Poriya, A., Poojary, D.: Type of NoSQL databases and its comparison with relational databases. *Int. J. Appl. Inf. Syst.* (2013)
18. Zablocki, J.: *Couchbase essentials*. Packt Publishing Ltd. (2015)

Security Analysis in the Architecture of the ATM Service



Gissela Quiña, Daisy Imbaquingo Esparza, Tatyana Saltos-Echeverria, José Jácome León, and MacArthur Ortega

Abstract In this research, a GAP analysis of the ATM service architecture was carried out; as a starting point, the current architecture of the ATM service was analyzed, identifying each of its components that make it up, to determine which components are suggested to be maintained, modified, and eliminated, to reduce the points of failure in the ATM system. In addition, the architecture of the ATM service was analyzed through a technical security procedure based on the good practices of the international Payment Card Industry Data Security Standard (PCI DSS) regulation. For the present study, a documentary and field research was carried out with data collection techniques such as the structured interview in which the Delphi method was used with 2 iterations, and this method was carried out to obtain efficient communication by the interviewer and the interviewee. In the first iteration, certified experts in computer security were interviewed, and in the second iteration, experts certified in the international PCI DSS standard were interviewed. The instrument used was a checklist, where compliance with the PCI DSS controls focused on the security of the ATM architecture service was evaluated. The results showed that thanks to the implementation of the GAP analysis, it was possible to determine improvements in terms of the diagram that makes up the architecture of the ATM service. In addition, through the technical procedure of the security analysis, the vulnerability of the medium and informative risk level was found. Finally, thanks to the implementation of the checklist and the interview, relevant information on the technological security of an automated teller machine (ATM) was obtained.

G. Quiña · D. I. Esparza (✉) · T. Saltos-Echeverria · J. J. León · M. Ortega
Universidad Técnica del Norte, Ibarra, Ecuador
e-mail: deimbaquingo@utn.edu.ec

G. Quiña
e-mail: gmquinap@utn.edu.ec

T. Saltos-Echeverria
e-mail: ksaltos@utn.edu.ec

J. J. León
e-mail: jgjacome@utn.edu.ec

M. Ortega
e-mail: mc.ortega@utn.edu.ec

Keywords Computer security of ATMs · PCI DSS · Computer audit at ATMs · GAP analysis

1 Introduction

The increase and evolution of technology means that society the need to carry out an analysis of threats, vulnerabilities, and risks to the computer assets are exposed, for that reason computer security has a very important role within organizations since they can suffer from some type of data corruption since in most organizations the data is stored by a computer system.

To minimize types of criminal computer activities within companies, there are security audits that review equipment and procedures to ensure the safety and security of the most important asset owned by the company, the information [9]. In addition, it is important to carry out an initial diagnosis of the audited company, through a GAP analysis, which is a measurement method to determine the gap between the performance of a variable and the expectations that are expected for that variable [12].

In the case of financial institutions, they present various attacks on ATMs, which used to be only physical attacks, such as card theft. However, criminals now frequently use malicious software to remove unauthorized cash from these computers [10].

Within logical attacks, security criminals violate the ATM network to control its server, where they install malware with the help of improved code, thus obtaining access to the internal command of the ATM and making the server order several endpoints that are infected to distribute cash [2].

Given the problems described above in ATMs, this research aims to carry out an initial security analysis in the architecture of the ATM service to identify, verify, and evaluate the vulnerabilities existing in them. In addition, information on the security of an ATM architecture was collected through a structured interview and a checklist to assess compliance with the international PCI DSS regulations. The regulation is focused on identifying security guidelines in ATMs worldwide. This is done because ATMs are equipment that stores and dispenses money; therefore, they are a point of attraction for attacks by cybercriminals.

2 Materials and Methods

For the development of this research on the security analysis in the architecture of the ATM service, activities were carried out to collect information and activities that allow the security analysis to be carried out in a technical and structured way.

Table 1 Reviewed articles

Informatic security			Computer security in ATMs		
Repository	Number of items	Item year	Repository	Number of items	Item year
IEEE			IEEE	5	2017–2020
SCOPUS	1	2019	SCOPUS	1	2019
Computer audit in ATMs					
Repository		Number of items		Item year	
IEEE		1		2016	
SCOPUS		3		2018 - 2019	

Source Own elaboration

2.1 Documentary Research

To obtain information on the issue raised, dentary research was carried out in high-impact indexed journals such as IEEE and SCOPUS using search strings such as:

TITLE-ABS-KEY (seguridad AND informática) AND PUBYEAR > 2016 AND PUBYEAR < 2021

TITLE-ABS-KEY (computer AND security AND in AND atms) AND PUBYEAR > 2016 AND PUBYEAR < 2021.

The articles that were reviewed are related to computer security in a general context, computer security in ATMs, and computer auditing in ATMs, taking into account the years 2016–2021, giving a total of 11 articles reviewed, which were relevant to the investigated topic.

Table 1 shows the bibliographic review carried out.

2.2 Field Research

A field investigation was carried out by implementing a technical procedure to analyze existing vulnerabilities in ATMs.

As an initial step, a GAP analysis is performed to evaluate the current state of the ATM service architecture and, based on this, propose points for improvement, following the phases that are within the analysis such as current situation, proposed situation, and gaps or improvements.

In addition, a checklist was carried out to analyze the architecture of the ATM service to measure compliance with the international PCI DSS regulations.

Finally, a structured interview was conducted with certified experts in computer security and certified experts in the international PCI DSS regulations, to identify the existing vulnerabilities in ATMs and the positive and negative aspects involved in implementing the international PCI DSS regulations in ATMs.

2.2.1 Implementation of GAP or GAP Analysis to the ATM Service Architecture

To carry out the GAP analysis, an unreal name of the financial entity will be taken, which will be named “Cooperativa X” for confidentiality reasons; however, the data obtained from the financial institution is real.

Figure 1 shows the conceptual diagram of the current architecture of the ATM service with its respective elements that compose it, and the conceptual diagram belongs to Cooperativa X.

Table 2 describes each of the elements that make up the ATM service architecture diagram.

Then perform the analysis of the current architecture of the ATM service of Cooperativa X, with the information provided by the financial institution.

Figure 2 presents the conceptual diagram of the proposed situation of the ATM service architecture.

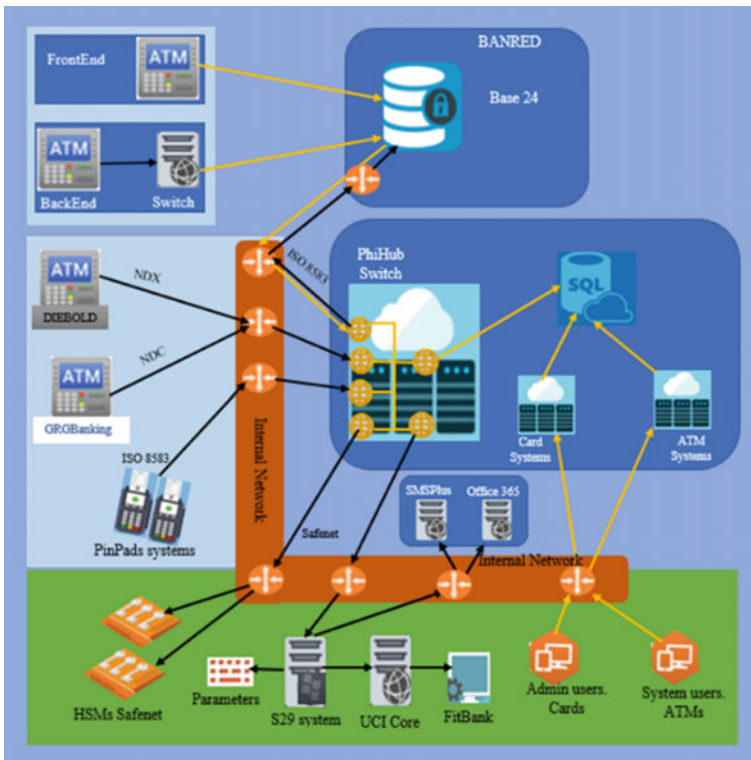


Fig. 1 Current architecture of the ATM service

Table 2 Elements of the conceptual diagram of the ATM architecture

Elements of ATM architecture	Description
Diebold terminals	Diebold brand ATMs recyclers and dispensers. They have the Agilis system installed to manage their devices and to manage the connection with the transactional switch emulating NDC (ATM protocol) messaging
Terminals GRGBanking	GRGBanking brand ATM dispensers. They have the YDC system installed to manage their devices and to handle the connection with the transactional switch, and they emulate NDC messaging
Terminals PINPADs	Generate or reset the passwords of the cardholders
PhiHub Switch	Responsible for the orchestration of transactions and the interconnection of the terminals
Card system	Life cycle management system for customer debit cards
ATM system	ATM administration system and transactional reports
SQL database	Systems data and configurations reside
SQL database	Carry out cryptographic processes such as generation of communication keys for terminals, PIN validation
Parameters database	Stores ISO8583 messaging settings and mappings
S29 System	Evaluates whether the transactions were approved by the Core FitBank to send SMS or email notifications in cash withdrawal and/or password change transactions
UCI Core	Entry system or interface to the Core FitBank
FitBank Database	Financial core
SMSPlus	SMS notification system
Office 365	Email notification system
Terminals FrontEnd	ATMs of other financial entities managed by BANRED's Base 24 platform
Terminals BackEnd	ATMs of other financial entities are managed by the systems of each financial entity
Switch BackEnd	Transactional switch of other financial entities responsible for the orchestration of transactions and the interconnection of their terminals with the BackEnd Switch and with BANRED's Base 24 platform
Base 24—BANRED	BANRED transactional switch is responsible for managing FrontEnd terminals and the orchestration of interbank transactions

Source Own elaboration

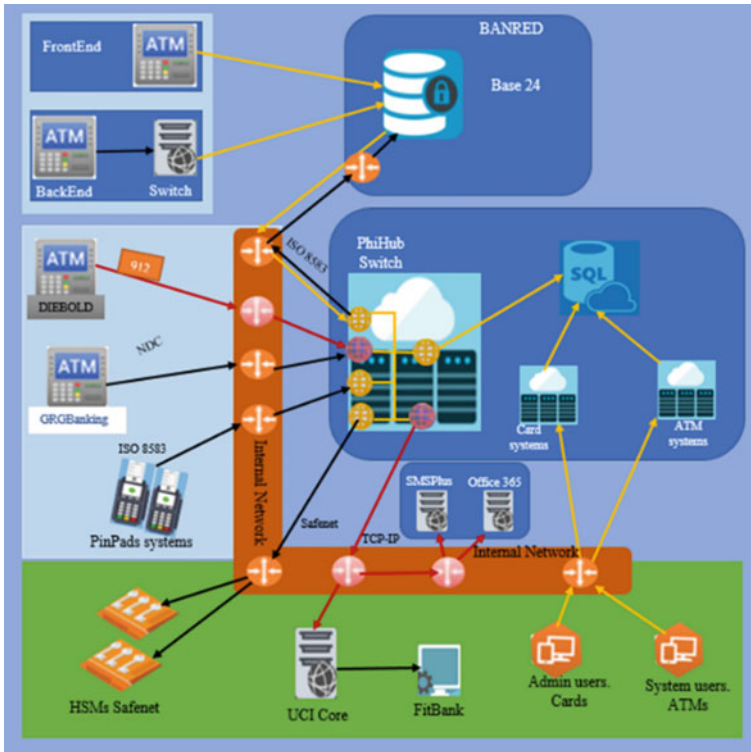


Fig. 2 Proposed architecture of the ATM service

2.2.2 Technical Procedure for Security Analysis in Automated Teller Machine Architecture

Each of the points that are part of the security analysis are described below.

Definition of Tools

Table 3 lists the tools used to perform the security analysis in the ATM architecture.

Attack Scenario

Table 4 presents the information of each ATM where a vulnerability analysis was carried out, using the tools described in Table 3.

Table 3 Tools to audit

Tool name	Description
Nmap	Nmap is a very flexible tool that performs security tests, and it has a scripting engine and an extensive and customizable scripting library, allowing a wide set of advanced and specialized security tests in a wide range of scenarios [13]
Nessus	The Nessus tool is used to perform vulnerability scans, obtaining a list of these [1]
Kali Linux	Kali Linux is the most popular software package for penetration testing and security auditing, it consists of hundreds of pre-designed tools, the tools are divided into sections, and each section performs different tasks with the same objective, to do penetration tests [6]

Source Own elaboration

Table 4 Devices to analyze

Cashier's name	Cashier brand
Cashier 1	GRG
Cashier 2	Diebold

Source Own elaboration

Vulnerability Detection, Verification, and Evaluation

In recent times, it has been observed that the security measures implemented in ATMs to access funds smoothly and safely have many loopholes that are exploited by hackers [14].

To detect, verify, and evaluate the vulnerabilities, each of the tools described above was used. Using the Nmap tool, the search for IP addresses and open ports of each ATM analyzed was carried out. Likewise, the Nessus tool was used to scan the existing vulnerabilities in ATMs, and this tool shows a complete report of each vulnerability found, with its respective level of impact and the recommendation that should be followed to mitigate each of the vulnerabilities.

Table 5 shows each of the vulnerabilities found within the analysis of the GRG brand ATM.

Table 6 shows each of the vulnerabilities found within the analysis of the Diebold brand ATM.

2.2.3 Collection of Information Through a Checklist Applying the International PCI DSS Regulations

The international PCI DSS regulation implies the implementation of technological and operational controls within financial companies. Organizations must implement technological controls with computer security tools to protect their infrastructure, through the LAN network, the wireless network, IT assets, databases, and Web applications [3].

Table 5 Vulnerabilities of the GRG brand ATM

Vulnerability	Impact	Detail/recommendation
IP forwarding enabled	Means	The remote host has IP forwarding enabled, which an attacker can exploit to route packets through the host. Unless the remote host is a router, it is recommended that you disable IP forwarding
Addresses MAC Ethernet	Informative	Este plugin reúne direcciones MAC de varias fuentes y consolida en una list
Host fully qualified domain name (FQDN) resolution	Informative	It is possible to resolve the name of the remote host
Nessus Scan Information	Informative	This plugin shows, for each tested host, information about the scan: edition of scanner used, scanner type, the port scanner used, range of scanned ports, scan date, scan duration

Source Own elaboration

Table 6 Vulnerabilities of the Diebold brand ATM

Vulnerability	Impact	Detail/recommendation
Addresses MAC Ethernet	Informative	This plugin gathers MAC addresses from various sources and consolidates them into a list
Microsoft Windows SMB2 dialects supported (remote check)	Informative	Nessus was able to get the set of SMB2 dialects running on the remote host by sending an authentication request to port 139 or 445
Nessus Scan Information	Informative	This plugin shows, for each host, tested, information about the scan: scanner edition (Nessus or Nessus Home), Nessus engine version, scan policy used, scanner IP, scan date, scan duration
Windows NetBIOS/SMB Remote Host Information Disclosure	Informative	The remote host is listening on UDP port 137 or TCP port 445 and responding to NetBIOS or SMB requests

Source Own elaboration

A checklist was prepared to measure compliance with the requirements and controls of the international PCI DSS regulations. The requirements and controls that apply only to the security of the ATM service architecture were chosen. This process was carried out with experts certified in the international PCI DSS regulations.

Table 7 describes the requirements of the international PCI DSS standard that were applied to assess compliance and verify the security of ATMs.

Table 7 Compliance with the requirements of the international standard PCI DSS

PCI DSS requirements and controls	Achieve
Requirement 2: Do not use vendor-supplied defaults for system passwords and other security parameters	
Always change vendor defaults and remove or disable unnecessary default accounts before installing a system on the network	Yes
Develop configuration standards for all system components. Make sure these standards address all known security vulnerabilities and are consistent with industry-accepted high system security standards	Yes
Implement only one primary role per server to prevent roles that require different levels of security from coexisting on the same server (e.g., Web servers, database servers, and DNS must be deployed on separate servers)	No
Enable only the necessary services, protocols, daemons, etc., as required by the system role	No
Implement additional security features for required services, protocols, or daemons that are not considered secure	Yes
Configure system security settings to prevent misuse	Yes
Encrypt all non-console administrative access using strong encryption	No
Keep an inventory of system components that are within the scope of PCI DSS	Yes
Ensure that security policies and operating procedures for managing vendor default settings and other security settings are documented, implemented, and known to all affected parties	Yes
Requirement 5: Protect all systems against malware and update antivirus software or programs regularly	
Deploy antivirus software on all systems which are typically affected by malicious software (especially personal computers and servers)	No
Ensure that antivirus programs can detect and remove all known types of malicious software and protect systems against them	Yes
For those systems that are not typically affected by malicious software, conduct regular assessments to identify and evaluate the malware threats that may appear to determine whether or not antivirus software is required to deploy on those systems	Yes
Ensure the antivirus mechanisms are up to date	Yes
Ensure that antivirus mechanisms are working actively and that users cannot disable or alter them unless specifically authorized by management in particular cases and for a limited time	Yes
Ensure that the security policies and operating procedures that protect the systems are documented, implemented, and known to all affected parties	Yes
Requirement 6: Develop and maintain secure systems and applications	
Establish a process to identify security vulnerabilities using known external sources to obtain information about security vulnerabilities, and assign a risk rating (e.g., “high,” “medium,” or “low”) to security vulnerabilities recently discovered security	Yes
Ensure all software and system components have vendor-supplied security patches installed that protect against known vulnerabilities. Install security patches within one month of their release	No

(continued)

Table 7 (continued)

PCI DSS requirements and controls	Achieve
Removal of data and accounts from system components before production systems are activated	Yes
Requirement 7: Change control procedures should include the following:	
Incident documentation	Yes
Documented change approval by authorized parties	Yes
At the end of a significant change, all relevant PCI DSS requirements must be implemented on all new or modified systems and networks, and documentation updated as applicable	Yes
Ensure that security policies and operating procedures for developing and maintaining secure systems and applications are documented, implemented, and known to all affected parties	Yes
Requirement 8: Identify and authenticate access to system components	
Give all users a unique ID before allowing them access to system components	Yes
Immediately terminate access to any dismissed user	Yes
Limit repeated login attempts by locking the user ID	Yes
Make all authentication credentials (such as passwords/phrases) unreadable during transmission and storage on all system components using strong cryptography	No
Verify the identity of a user before modifying any authentication credentials, for example, reset the password, deliver new tokens, or generate new keys	Yes
Do not allow a person to submit a new password/phrase that is the same as any of the last four passwords/phrases used	Yes
Document and communicate authentication policies and procedures to all users, including the following: <ul style="list-style-type: none"> • Guidelines on how users should protect authentication credentials • Instructions not to select previously used passwords • Instructions for changing passwords if the password is suspected to be at risk 	Yes
Authentication mechanisms must be assigned to a single account and not shared among multiple accounts	Yes
Ensure that security policies and identification and authentication operating procedures are documented, implemented, and known to all affected parties	Yes
Requirement 10: Track and monitor all access to network resources and cardholder data	
Implement audit trails to link all access to system components with specific users	Yes
Implement automatic audit trails on all system components to reconstruct the following events:	
Verify that all actions performed by people with administrative privileges are logged	Yes
Verify that access to all audit trails is logged	Yes
Verify that all changes, additions, and deletions for an account with administrative or root privileges are logged	Yes
Verify that the following is logged <ul style="list-style-type: none"> • Initialization of audit records • Stop or pause audit logs 	Yes

(continued)

Table 7 (continued)

PCI DSS requirements and controls	Achieve
Recording at least the following entries of the audit system component trail for each event:	
Identificación de usuarios	Yes
Event type	Yes
Date and time	Yes
Indication of success or failure	Yes
Origin of the event	Yes
Identity or name of the affected data, system components, or resources	Yes
Protect audit trails so that they are not modified	Yes
Retain audit trail history for at least one year, with a minimum availability for analysis of three months (e.g., online, archived, or retrievable for backup)	Yes
Requirement 11: Regularly test security systems and processes	
Implement processes to determine the presence of wireless access points, and detect and identify all authorized and unauthorized wireless access points	Yes
Perform internal and external network vulnerability scans at least quarterly and after every significant network change	No
Use intrusion detection and intrusion prevention techniques to detect or prevent network intrusions	Yes
Ensure that security policy and operational procedures for monitoring and verifying security are documented, implemented, and known to all affected parties	Yes
Requirement 12: Maintain a policy that addresses information security for all personnel	
Establish, publish, maintain, and distribute a security policy	Yes
Review the security policy at least once a year and update it when changes are made to the environment	Yes
Develop usage policies for critical technologies and define how to use them correctly	Yes
Ensure that security policies and procedures clearly define the information security responsibilities of all personnel	Yes
Establish, document, and distribute security policies and procedures	Yes
Monitor and analyze alerts and security information and communicate them to the appropriate personnel	Yes
Establish, document, and distribute security incident response and escalation procedures to ensure timely and effective handling of all situations	Yes
Train staff immediately upon hire and at least once a year	Yes
Require staff to make a statement at least once a year that they have read and understood the company’s security policy and procedures	Yes
Screen potential staff prior to hiring to minimize the risk of attacks from internal sources	Yes
Implement an incident response plan. Prepare to respond immediately to a system failure	Yes

Source Own elaboration

2.2.4 Collection of Information Through a Structured Interview Applying the Delphi Method

The interview is a technique in which one person requests information from another about a specific problem. The structured interview is one that takes place within a more rigid framework, where the interviewer uses a format of questions that are not changed or disrespected in their order or meaning [5].

The Delphi method is a consensus tool that provides excellent reliability when applying qualitative criteria. This method must have a logical order according to the goals achieved, and the problem to be analyzed must be clear to develop the precise instruments with experts on the subject [8].

Using the method Delphi data collection was performed in two iterations, with structured interviews, and complying with the phases that are within this method, such as the definition phase and expert selection phase. The first iteration was performed by certified experts in the security of the ATM service architecture, and the second iteration was performed by experts certified in the international PCI DSS standard. En la Table 8 se presenta las fases del método Delphi con su respectiva descripción aplicando dos iteraciones.

Table 8 Phases of the Delphi method

Phases	Description
Definition phase 1	Identify existing vulnerabilities in the architecture of the ATM service through the implementation of the international PCI DSS standard to mitigate the risks to which these devices are exposed
Phase 2 of selection of experts	The interview was conducted with certified specialists in the security of the architecture of the ATM service and specialists certified in the international standard PCI DSS
Phase 3 of preparation of consultations	In the first iteration, a structured interview was developed, directed to the personnel in charge of the security of the ATM service In the second iteration, a structured interview was developed, directed to specialists in the security of the ATM service and experts certified in the international regulations PCI DSS and PCI PTS
Results phase 4	The result of the interview was carried out in each iteration

Source Own elaboration

3 Results

Thanks to the literature review in the bibliographic databases such as IEEE and SCOPUS, it was possible to obtain information that was relevant to the subject of this present investigation; with a total of 11 articles, 6 articles were found in the IEEE bibliographic base and 5 articles were found in the SCOPUS bibliographic database.

Once the proposed architecture of the ATM service has been analyzed, the GAP analysis is carried out, to specify the components that must be modified and eliminated. Table 9 shows the GAP analysis with the components of the diagram that must be modified or eliminated.

Using software tools to assess existing vulnerabilities in the ATM service architecture, vulnerabilities were found in the IP network analysis of two GRG-branded ATMs that were detected, verified and evaluated under the DIEBOLD brand.

Table 10 shows the total vulnerabilities found in the GRG brand ATM and the level of risk to which each vulnerability belongs.

Figure 3 presents the summary of vulnerabilities that were found in the analyzed ATM of the GRG brand, 1 medium-level vulnerability was found representing 25%, 3 informational vulnerabilities were found representing 75%, and no level vulnerabilities were found critical, high, and low.

Table 11 shows the vulnerabilities found in the Diebold brand ATM and the level of risk to which each vulnerability belongs.

Table 9 GAP analysis

Elements	Description	Action
Terminals Diebold	It suggests installing the version of Agilis that emulates messaging 912 to take advantage of all the functionalities for handling screens with animations	Modify
S29 System Parameter database	The S29 System in each transaction accesses its parameter database to retrieve the configuration of the frame sent by the transactional switch and the setting of the UCI Core system frame and its respective mapping. This functionality can be executed by the transactional switch, as happens in other financial entities that have FitBank as a core financier	Remove the 2 components
PhiHub Switch	Create an interface or direct connection with the UCI Core component to connect with the FitBank authorizing system Assesses if the transactions were approved by Core FitBank to connect with the SMSPlus and Office 365 notification systems to send SMS or email notifications in cash withdrawal and/or password change transactions	Modify

Source Own elaboration

Table 10 Total vulnerabilities of the GRG ATM

Vulnerability name	Critical	High	Medium	Low	Informative
IP forwarding enabled	0	0	1	0	0
Ethernet MAC addresses	0	0	0	0	1
Host fully qualified domain name (FQDN) resolution	0	0	0	0	1
Nessus Scan Information	0	0	0	0	1

Source Own elaboration

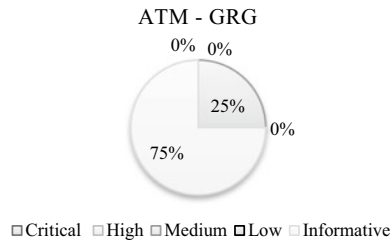


Fig. 3 Total vulnerabilities of the GRG brand ATM

Table 11 Total vulnerabilities of the Diebold ATM

Vulnerability name	Critical	High	Medium	Low	Informational
Ethernet MAC addresses	0	0	0	0	1
Microsoft Windows SMB2 Dialects Supported (Remote Check)	0	0	0	0	1
Nessus Scan Information	0	0	0	0	1
Windows NetBIOS/SMB Remote Host Information Disclosure	0	0	0	0	1

Source Own elaboration

Figure 4 presents the summary of vulnerabilities found in the analyzed Diebold brand ATM; 4 informational vulnerabilities were found representing 100%.

Table 12 shows the results of compliance with the requirements and controls of the international PCI DSS regulations, which were applied to the analyzed ATMs.

Figure 5 shows the 59 controls are established in the international PCI DSS regulations, of which they comply with 52 controls representing 88% and do not comply with 7 controls representing 12%.

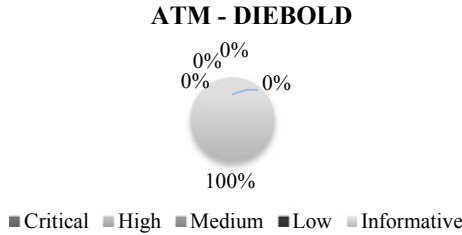


Fig. 4 Total of vulnerabilities of the GRG brand ATM

Table 12 Result of compliance with the international regulations PCI DSS

PCI DSS requirement	Total, controls	Yes	No
Requirement 2: Do not use vendor-supplied defaults for system passwords and other security parameters	9	6	3
Requirement 5: Protect all systems against malware and update antivirus software or programs regularly	6	5	1
Requirement 6: Develop and maintain secure systems and applications	7	6	1
Requirement 8: Identify and authenticate access to system components	9	8	1
Requirement 10: Track and monitor all access to network resources and cardholder data	13	13	0
Requirement 11: Regularly test security systems and processes	4	3	1
Requirement 12: Maintain a policy that addresses information security for all personnel	11	11	0

Source: Own elaboration

Fig. 5 Compliance with international PCI DSS regulations



Table 13 presents the result of the interview applied in the first iteration using the Delphi method.

The personnel interviewed in the first iteration answered with clarity and precision each question asked; however, another iteration was carried out where the interview reformulated and to be able to capture positive and negative aspects that the implementation of the international PCI DSS regulations entails.

Table 14 shows the result of the interview applied in the second iteration using the Delphi method.

Table 13 Interview of the first iteration

Delphi method—first iteration	
Question	Answer
¿ What are the services that banks or cooperatives offer through ATMs?	Experts responded that the services offered by financial entities are: withdrawal of money, payments, and transfers
¿ What is your opinion regarding the quality of service offered by ATMs? ¿ Do you consider that the different users are satisfied?	Experts agree that some customers who use ATMs are not satisfied, as there are problems due to the lack of availability of the service they offer
¿ How do you think the security of the ATM service is currently?	Experts responded that the security of the ATM service is stable
¿ What situations of unavailability in the ATM service have you witnessed?	Experts agree that they have witnessed situations of unavailability, such as: out of service
¿ What flaws do you most often find in the security of ATM service?	Experts respond that they have witnessed failures in the security of the ATM service, such as lack of connection
¿ How do you identify the security flaws that are present in the logical infrastructure of an ATM?	Experts responded that they identify security flaws in the logical infrastructure through a monitoring system
¿ What solution would you give to improve these types of failures?	Experts responded that the solution is: toughen computer platforms and network traffic analysis
¿ How do you prepare to successfully face the new threats that appear in the ATM system?	Experts agree about using an adequate installation of computer security components that must be carried out in ATMs, such as antivirus, to face new threats
¿ Do you document the security policies to keep the systems safe and known to all affected parties?	Experts do document security policies so that they are known to all affected parties
¿ Do the ATMs have the good practices established in the international PCI DSS standard??	Experts responded that in most financial institutions, they do, but only as good practices, but financial institutions do not reach a certification
Maintaining an information security policy is a goal of the PCI DSS; ¿ Do you implement an ATM system incident response plan?	Experts responded that it is convenient with the implementation of an incident response plan to mitigate the risks to which they may be exposed
Maintaining a vulnerability management program is a goal of the PCI DSS; do you implement antivirus software on all systems that are generally affected by malicious software in ATMs?	Experts agree to carry out a vulnerability analysis with the implementation of antivirus software

Source Own elaboration

Table 14 Interview of the second iteration

Delphi method—second iteration	
Question	Answer
¿ How do you think the security of the ATM service is currently?	Experts responded that the security of the ATM service is high, medium, and stable, while other experts responded that since 2012, with the entry of the certifications, the resolution of the banking board JB-2012–2148 has improved
¿ Do you think that most attacks on ATMs are due to security problems and because they have insecure software why?	Experts responded that most of the attacks are due to security problems since not many institutions considering the minimum security standards required for the operation of ATMs
¿ Do you consider that the operating system used by an ATM is a weak point against attacks by cybercriminals and why?	Experts agree that the operating system is a weak point against attacks, since they have not been completely updated in all ATMs, and some ATMs maintain an OS that no longer has technical support from the manufacturer, that is why they become obsolete with many security problems or bugs
¿ What are the benefits of implementing IT security in ATMs?	Experts responded that the implementation of computer security in ATMs helps mitigate different types of risks, raise security levels, and minimize security risks concerning confidentiality, integrity, and availability of information
¿ Do you consider that the international regulations PCI DSS and PCI PTS manage to determine certain guidelines to reduce vulnerabilities in ATMs and why?	Experts respond that the regulations are a standard for the treatment of ATM information that help to define some guidelines since the minimum requirements must be considered in security over electronic channels
¿ How do you implement the PCI DSS and PCI PTS regulations in ATMs?	Experts implement regulations and evaluate ATMs, considering processing offices, databases, people who have access to data, among others In addition, they implement the controls indicated in the standards at the hardware, software, process, procedures, and policies level; other experts responded that it should be initiated with a GAP analysis and then plan the implementation of the controls
¿ What are the benefits of implementing the PCI DSS and PCI PTS regulations in ATMs?	Experts responded that the benefits of the PCI DSS regulation and the PCI PTS regulation are obtained when they are adequately limited and implemented since they provide control measures, regulation of the media, and payment processes In addition, they respond that reducing information security risks is another benefit of regulations

(continued)

Table 14 (continued)

Delphi method—second iteration	
Question	Answer
¿What negatives have you witnessed regarding the implementation of PCI DSS at ATMs?	Experts responded that not all companies that carry out a real implementation of the controls are properly implemented, so this is considered a negative aspect in the implementation of the regulations

Source Own elaboration

In this second iteration, the interviewees answered each question asked, obtaining a better result regarding the implementation of the PCI DSS with the positive and negative aspects.

4 Discussion

To analyze the security in the architecture of the ATM service, the results reflected that by performing the GAP analysis on the conceptual diagram of the ATM service architecture, it was established to modify and eliminate components that do not generate value. The time of making a transaction and are points of failure of the system. Likewise, a technical procedure was carried out, where the existing vulnerabilities in GRG and Diebold brand ATMs were detected, verified, and evaluated, obtaining a total of 4 vulnerabilities; in each ATM, the majority of them were informative. In addition, a checklist was made where compliance with the requirements and controls of the international PCI DSS regulation was measured, evaluating 59 controls focused on the security of the technological infrastructure of an ATM, resulting in compliance with 52 controls that they represent 88% and do not comply with 7 controls giving a total of 12%; this means that the evaluated ATMs comply with most of the controls of the regulation; however, with the existence of unfulfilled controls, it is recommended to provide an immediate solution to obtain quality technological security. This research was carried out because at ATMs several logical attacks based on software and this is corroborated by Jacintha et al. [7] who highlights that digital transactions play a very important role in all factors of ATMs and have a higher priority in fraudulent ATMs, for that reason recommends using protective equipment against threats to which they are exposed. Likewise, Patoliya and Desai [11] highlight that financial users prefer to make cash withdrawals at an ATM, and for that reason, banks focus on the security of this equipment, adequately protecting them from criminal activities, continuing rules, and procedures for analyzing security. Embarak [4] describes that logical attacks on ATMs are done using external electronic devices or malicious software to violate the confidentiality and authenticity of ATM transactions. This method is affected by the ATM operating system, communication system, and software.

Yulianto et al. [15] recommend that, to comply with the international PCI DSS regulations, a simple, reliable, and easy-to-use evaluation mechanism or tool is needed to advance the general security status of an organization.

Once the thoughts of each aforementioned authors have been analyzed and the results of the present investigation have been analyzed, it can be seen that it is necessary to carry out an analysis of computer security within ATMs. Likewise, it is important to carry out a technical procedure to mitigate the risks to which ATMs are exposed since they are devices that are always under constant computer attacks.

5 Conclusions

Through the literature review, adequate information was obtained which was selected for the research topic in the IEEE and SCOPUS bibliographic databases, filtering articles on computer security in general, computer security in ATMs, and computer auditing in ATMs, varying from 2016 to 2021.

With the implementation of the GAP analysis in the architecture diagram of the ATM service, improvements were established: eliminating and modifying the components that are part of the points of failure of the system and thus achieve improved performance in terms of conducting transactions.

With the technical security procedure of the ATM service architecture using software tools, the vulnerabilities existing in this equipment were analyzed, where 4 vulnerabilities were found in the GRG brand ATM, 1 vulnerability belongs to at a medium risk level, and 3 vulnerabilities at the information risk level, while in the Diebold brand ATM, 4 vulnerabilities were found, all of the information risk level.

Through a checklist to comply with the controls of the international PCI DSS regulations, security guidelines were identified and evaluated in ATMs to protect hardware and software against attacks that could compromise the confidential data of each user of a financial organization.

Through a structured interview carried out with certified experts in computer security and experts certified in the international PCI DSS regulations, the necessary information was obtained to verify how vulnerable ATMs are to computer attacks by cybercriminals and what positive aspects and negative is the implementation of the international PCI DSS regulation in ATMs.

References

1. Alzahrani, M.E.: Auditing Albaha University network security using in-house developed penetration tool. *J. Phys. Conf. Ser.* **978**(1) (2018). <https://doi.org/10.1088/1742-6596/978/1/012093>
2. Ansar, S.A., Kumar, S., Khan, M.W., Yadav, A., Khan, R.A.: Enhancement of two-tier ATM security mechanism: towards providing a real-time solution for network issues. *Int. J. Adv. Comput. Sci. Appl.* **11**(7), 123–130 (2020). <https://doi.org/10.14569/IJACSA.2020.0110717>

3. Casallas, H., Valderrama, J., Vargas, J.: Guía de implementación de herramientas tecnológicas dirigida a las PYMES para dar cumplimiento a la norma internacional PCI DSS v3.0. 171 (2015)
4. Embarak, O.H.: A two-steps prevention model of ATM frauds communications. In: 2018 Fifth HCT Information Technology Trends (ITT), pp 306–311 (2018). <https://doi.org/10.1109/CTIT.2018.8649551>
5. Guerrero, M.: La investigación cualitativa (2016). <https://doi.org/10.33890/innova.v1.n2.2016.7>
6. Gunawan, T.S., Lim, M.K., Zulkurnain, N.F., Kartiwi, M.: On the review and setup of security audit using Kali Linux. Indonesian J. Electr. Eng. Comput. Sci. **11**(1), 51–59 (2018). <https://doi.org/10.11591/ijeecs.v11.i1.pp51-59>
7. Jacintha, V.M.E., Rani, S.J., Beula, J.G., Johnslly, J.J.: An extensive resolution of ATM security systems. In: 2017 Third International Conference on Science Technology Engineering Management (ICONSTEM), pp. 934–938. <https://doi.org/10.1109/ICONSTEM.2017.8261340>
8. Martínez-Ezquerro, J.D., Ruiz-Cejudo, S.M., Bustamante-Fuentes, A., Díaz-Badillo, Á., García-Oropesa, E.M., López-Sosa, E.B., Martínez-López, Y.E., Moctezuma-Chávez, Ó.O., Nava-González, E.J., Perales-Torres, A.L., Pérez-Navarro, L.M., Rosas-Díaz, M., López-Alvarenga, J.C., Grupo, S.: Expert consensus in times of COVID-19: Health applications of the Delphi method. *Cirugia y Cirujanos (English Edition)* **89**(1), 120–129 (2021). <https://doi.org/10.24875/CIRU.20000936>
9. Muñoz Hernández, H., Zapata Cantero, L.G., Requena Vidal, D.M., Ricardo Villadiego, L.: Computer risks and alternatives for computer security in accounting systems in Colombia. *Revista Venezolana de Gerencia* **24**(2), 528–541 (2019). <https://doi.org/10.37960/REVISTA.V24I2.31508>
10. Ogata, H., Ishikawa, T., Miyamoto, N., Matsumoto, T.: An ATM security measure for smart card transactions to prevent unauthorized cash withdrawal. *IEICE Trans Inf Syst* **E102D**(3), 559–567 (2019). <https://doi.org/10.1587/transinf.2018EDP7136>
11. Patoliya, J.J., Desai, M.M.: Face detection based ATM security system using embedded Linux platform. In: 2017 2nd International Conference for Convergence in Technology (I2CT), pp. 74–78 (2017). <https://doi.org/10.1109/I2CT.2017.8226097>
12. Puspitasari, D., Kholdani, A.F.: Penerapan sistem Pendukung keputusan dengan analisis gap untuk pemilihan mahasiswa berprestasi tingkat fakultas (pada fakultas teknologi informasi uniska). *Technologia* **8** (2017). <https://doi.org/10.31602/tji.v8i1.716>
13. Redondo, J.M., Cuesta, D.: Towards improving productivity in NMAP security audits. *J. Web Eng.* **18**(7), 539–578 (2019). <https://doi.org/10.13052/jwe1540-9589.1871>
14. Swathi, H., Joshi, S., Kumar, M.K.K.: A novel ATM security system using a user defined personal identification number with the aid of GSM technology. In: 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAIECC), pp. 1–5 (2018). <https://doi.org/10.1109/ICAIECC.2018.8479533>
15. Yulianto, S., Lim, C., Soewito, B.: Information security maturity model: a best practice driven approach to PCI DSS compliance. In: 2016 IEEE Region 10 Symposium (TENSYMP), pp. 65–70 (2016). <https://doi.org/10.1109/TENCONSpring.2016.7519379>

Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution



Filomena Castro Lopes , Paula Morais , and Abílio Cardoso

Abstract The benefits recognized by integrated systems, integration, and information quality are beginning to prove insufficient to support decision-making and organization competitiveness. In this context, there is a marked growth of business intelligence (BI) systems, recognizing their analytical capacity and data integration. However, BI implementation projects bring to light two relevant technological critical success factors (CSFs): systems integration and information quality. This article reports the implementation of a BI project in a higher education institution. This project corroborates the literature highlighting the relevance of the technological CSF in achieving a successful BI implementation. Some recommendations to overcome difficulties regarding the technological CSF are also proposed.

Keywords Business intelligence · Decision support systems · Information systems integration · Data quality · HEI data analysis

1 Introduction

The need for organizations to acquire competitive advantage through available information has become increasingly pressing. To be able to make informed decisions, organizations require quality information supported by technological information systems, with decision support systems being particularly relevant.

The original version of this chapter was revised: The author name has been corrected and the missing ORCID identifier has been updated. The correction to this chapter is available at https://doi.org/10.1007/978-981-16-9268-0_61

F. C. Lopes (✉) · P. Morais · A. Cardoso
REMIT—Research On Economics, Management, and Information Technologies, Universidade Portucalense, Dr. António Bernardino de Almeida 541, 4200-072 Porto, Portugal
e-mail: flopes@upt.pt

P. Morais
e-mail: pmorais@upt.pt

A. Cardoso
e-mail: abilic@upt.pt

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022, corrected publication 2022

J. L. Reis et al. (eds.), *Marketing and Smart Technologies*, Smart Innovation, Systems and Technologies 279, https://doi.org/10.1007/978-981-16-9268-0_14

Over the years, organizations have invested in information systems and technologies to better manage the information they need. Organizations moved from “monolithic” systems to systems integration and after that to integrated systems, characterized by centralization and information quality [1, 2].

Integrated systems brought benefits to organizations [2, 3] such as the integration of business processes, ensuring the flow of information between different areas, increased productivity, and improved access to quality information. In this context, the integration of information becomes crucial, on the assumption that the right repositories can be accessed and that the content is trustworthy, timely, and structured. However, in recent years, the ever-changing and more complex environment have led to business intelligence (BI) systems gaining relevance in organizations [4]. These systems are seen as an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance [5]. BI systems manage and gather the information stored in the various systems of the organization to monitor its performance indicators.

Many works have focused on identifying the factors that, if not carried out properly, can compromise the successful implementation of information systems in organizations. These factors are commonly referred to as critical success factors (CSFs). CSFs are a relevant issue, considering the importance and impact of information systems (ISs) in organizations, so these studies have been carried out for various types of information systems.

Nowadays, Higher Education Institutions (HEIs) are also operating in this dynamic and complex environment, facing a strong competition among themselves. Therefore, they are also implementing advanced analytical technologies, such as BI tools for analyzing their information and for supporting their internal quality assurance system (IQAS).

This article describes the implementation of a BI system in a HEI, with the aim of evaluating the critical factors that contributed to its successful implementation, having as reference the technological CSF identified in the literature.

This article is divided into 5 main sections: In this first section, the work is contextualized, justifying its relevance. Section 2 presents a brief literature review on the main concepts. The third section describes the BI implementation project, focusing on the technological CSF. The results are discussed in Sect. 4 and Sect. 5 concludes.

2 Literature Review

2.1 Critical Success Factors in BI Implementations

There are several CSF identified in the literature, common to different types of information systems, such as top management support, alignment with the organization’s strategy, business knowledge, user involvement, and training of end users [6].

However, regarding the implementation of BI projects, technological factors appear that are not found for other types of systems [7], namely the integration between the BI system and other systems and the quality of data [8–11].

These CSFs are justified for a BI system since it is necessary to ensure the integration of information spread throughout the organization, whether in legacy systems or in enterprise systems (e.g., ERP, CRM), or in social systems and in many others. This integration of the BI system with the other systems is necessary to aggregate the different sources of information in the organization, thus enabling complex analysis of the organization's data and supporting the decision-making process. In turn, the integration will make it possible to unify the databases of the organization's different information systems. Data quality, sometimes also referred to as information quality, is the other technological CSF assigned to BI systems implementations. Several works [9, 11–14] suggest a set of measurements of information characteristics that evaluate information quality. A detailed analysis of those has highlighted four common characteristics for information quality: accuracy, completeness, timeliness, and consistency.

2.2 BI in Higher Education Institutions

The adoption of a quality assurance system (QAS) is, according to ISO 9001, a strategical decision of organizations. In recent years, HEIs have also implemented this kind of systems, for responding to requirements from assessment agencies, but also by recognizing that they are a tool for continuous improvement. In Europe, when speaking about QAS, the main point of reference is the standard and guidelines for quality assurance in the European Higher Education area document, devised by the European Association for Quality Assurance in Higher Education [15]. In Portugal, Portuguese Agency for Assessment and Accreditation of Higher Education (A3ES) also created a manual for the internal QAS audit process [16]. Both referred documents define a set of key process indicators (KPIs) that institutions should use.

In 2012, EDUCASE Center for applied research presented a report about the benefits, barriers, progress, and recommendations of the use of analytics in higher education [17]. At that time, the study concluded that analytics was viewed as important, but that data used at most institutions were still limited to reporting. Nowadays, to effectively monitor their indicators and support decision-making, many HEIs are already using BI systems.

According to some authors [18–20], one important advantage of implementing BI is the integration of the data from the different systems HEI use. HEIs face many challenges such as student retention, academic success, and costs reduction. They need to get a profound knowledge of their main stakeholders and processes to turn these difficulties into strengths.

In addition to support administrative and management activities, BI systems have been implemented to enhance one of the fundamental vectors of any institution's mission, education, as the examples below show.

Villegas-Ch et al. [21] say that most of the HEI systems store large amounts of data about students socioeconomic and academic variables, but do not use these data to generate knowledge about their students; therefore, they propose a BI framework that can be used to influence the development of learning.

Jayakody and Perera [22], considering the importance of identifying, guiding, and monitoring less-able students to develop them toward their full potential, describe the development of a BI framework to identify and compare less-able and able students.

Kabakchieva [18] describes an experience of using BI for analyzing the student performance with the aim of improving the educational process.

3 Case Description

This section describes the implementation of a BI system at Universidade Portucalense, with the objective of evaluating the critical factors that contributed to the success of its implementation, having as reference the two technological CSF identified in the literature.

The BI implementation process went through several stages: (i) analysis of the business processes, as well as of the HEI's mission and strategy; (ii) analysis of the pre-existing indicators for each process; (iii) literature review to identify and define the KPI; (iv) KPI validation by the members of staff responsible for each process and ensuing adjustments to the list of KPI resulting from the validation; (v) identification of the information systems implemented in the institution; (vi) analysis of the information systems to identify the necessary information sources for the KPI; (vii) definition of the users access profile, and (viii) selection and implementation of BI technological solution. In this project, Qlick Sense was the BI tool used [23].

In the context of this work, the relevant stages are stages v, vi, and viii.

3.1 HEI Characterization

Universidade Portucalense (UPT) is a private HEI founded in June of 1986, located in Porto (Portugal). UPT is organized into six education departments and 4 research centers. Its main scientific areas are Law; Economics; Management; Information Technology; Heritage and Tourism; Psychology and Education; Architecture and Multimedia. Today, UPT has 3300 students. As a private institution, its main governance structure is divided into a Board and a Rectory.

The university has unambiguously assumed the principle that quality and quality assurance are one of the fundamental vectors for its operation and development.

Having its continuous improvement in mind, the university has implemented an internal quality assurance system (IQAS).

The IQAS includes, among others, the monitoring of processes as a way of continuously improving UPT's activities. UPT has identified its processes and has established methods for their operation, control, and monitoring, which are echoed in the process map contained in the quality management manual and in the performance monitoring plan which already includes a set of KPI. The processes are the matrix of the structure of the IQAS, and its performance measure is decisive for UPT quality assurance.

The process map was built following the Porter [24] idea of the value chain and is divided into primary and support processes. There are nine primary processes: marketing and communication (MC), course design and development (CDD), teaching management (TM), innovation, research and development (IRD), application management (AM), enrollment management (ER), inter-institutional and community collaboration (ICC), internationalization (I), and skills and careers management (SCM) and nine support processes: supply management (SM), scholarship management (ScM), information resource management (IRM), administrative and financial management (AFM), infrastructure management (IM), information system management (ISM), strategic planning (SP), human resource management (HRM), and quality assurance management (QAM).

3.2 Information Systems Integration

The information needed to monitor HEI's performance indicators is spread across several information systems, some internal (24) and others external (5). Figure 1 depicts the UPT information systems portfolio highlighting the systems that support the 18 UPT processes.

The BI system will have to integrate these systems to respond to the 380 KPI defined in this project. During stage "vi", it was noticed that 16 processes lack full technological support and that there are still many of those supported by Excel and Word files.

3.3 Data Quality

For each process and its KPI, the data quality issues considering the four common characteristics (consistency, completeness, accuracy, and timeliness) were identified as displayed in Table 1.

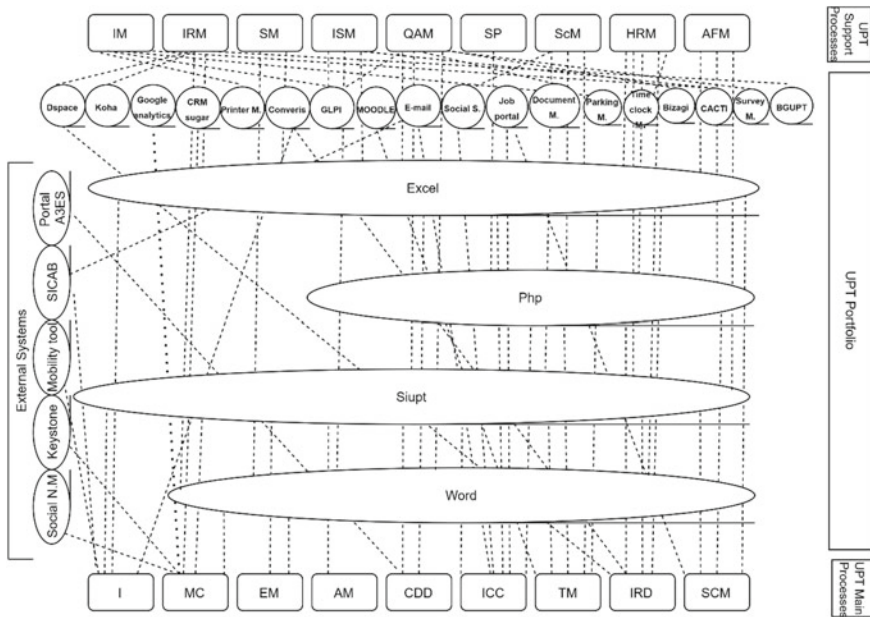


Fig. 1 Information systems portfolio supporting UPT processes

4 Discussion

The implementation of this BI system has proved to be advantageous. However, it made the institution aware of new needs since, as described in the previous sections, various problems have arisen in this implementation. Technological CSFs were detected, however, integration of systems was not a problem; although UPT has a vast IS portfolio supporting its processes, the BI tool used was able to integrate all the digitalized information coming from the different sources.

The lack of information system support to feed some KPI was a problem. The institution does not have a completely integrated IS, and many “monolithic” different systems using different data formats coexist. Moreover, some activities are still supported by personal Excel sheets and informal data sources such as Word and other have no computerized support. Obviously, this lack of information systems supports jeopardizes obtaining several KPI, diminishing the project’s success. So, this BI implementation allowed to alert the prioritization of new IS implementations.

As corroborated with the literature, data quality was reality a CSF. In this implementation, this factor is strong related to organizational factors such: user education and training and user participation, as also highlighted in the literature [9]. Table 1 enhances this problem mainly in two data characteristics. There is lack of consistency in the data needed to feed 40 KPI. Some of the problems are related to the lack of data typification (for example: extracurricular activities, student dropout reasons, curricular units grade ranges, inter-institutional collaboration events). Considering

Table 1 Data quality problems identified in KPI generation

Processes	KPI	Consistency	Completeness	Accuracy	Timeliness
Teaching management	60	1	9	4	
Course design and development	14	2	1		
Innovation, research, and development	33		4	2	3
Application management	14	1	2		1
Enrollment management	34	2	7		1
Marketing and communication	25	3			
Skills and careers management	10				1
Inter-institutional and community collaboration	9	7			
Internationalization	31	3	13	2	
Supply management	8		1	1	
Scholarship management	11	5	1		2
Information resource management	21	1	1		
Administrative and financial management	21	3			
Infrastructure management	21	8	2	1	
Information system management	12	3			
Human resource management	29	2	1		10
Strategic planning	9				
Quality assurance management	18	2	1		

completeness, the problem arises in 43 KPI. It was observed that the institution does not collect the necessary information to support some of the identified KPI. But there are also problems regarding accuracy (in 9 KPI) and timeliness (in 16 KPI). The issues concerning these two data problems are primarily related to human factors, mainly due to the lack of sensibility and responsibility regarding data introduction in the different systems.

4.1 Recommendations

A successfully BI project implementation is not simple. Institutions should define policies to minimize data problems, such as the typification of data, wherever possible. To successfully implement a BI project, it is necessary for CIOs to design more holistic policies regarding data quality. But the problem is not only at the CIO level. It will also be necessary to raise awareness and make users of the IS portfolio

responsible for the quality of the data. We can highlight that the IS social component is a crucial component in the success of these systems, reinforcing that the involvement and training of users are aspects that cannot be neglected.

In addition, only institutions with a high level of digitization will be able to achieve greater benefits with the implementation of BI projects, since the lack of information in digital support does not allow the integration of the information and consequently, the achievement of KPI.

5 Conclusions

As stated in the literature, BI implementation projects bring to light two relevant technological critical success factors: systems integration and information quality.

Although the tool used (Qlick) has integrated information from different sources, the lack of necessary information in digital support was felt. The problem detected has to do with the lack of IS support in several processes; the level of digitization is still relatively low in some of them, in the Innovation, research and development, teaching management, marketing and communication processes, internationalization, human resources management, IS management and infrastructure management. This analysis enabled the identification of IS development needs.

Information quality really stood out as a (missing) critical success factor in this implementation. This was caused by the existence of a lot of information in personal tools such as Word and Excel and IS developments at different times, responding to immediate needs. These facts led to the need to review the data design in an integrated manner.

References

1. Martins, V.: *Integração de Sistemas de Informação Perspetivas, Normas e Abordagens*. Edições Sílabo (2006)
2. Myerson, J.M.: *Enterprise Systems Integration*. CRC Press Company (2002)
3. Rotchanakitumnuai, S.: Measuring ERP implementation success with a balanced scorecard. In: *Accounting Information Systems (SIGASYS)* (2017)
4. Luftman, J.N., Ben-Zvi, T.: Key issues for IT executives 2010: judicious IT investments continue post-recession. *MISQ Exec.* **9**, 263–273 (2010)
5. Ong, L., Siew, P.H.: Empirical investigation on the evolution of BI maturity in Malaysian organizations. In: Tavana, M., Puranam, K. (eds.) *Handbook of Research on Organizational Transformations through Big Data Analytics*. IGI Global (2015). <https://doi.org/10.4018/978-1-4666-7272-7.ch012>
6. Fortune, J., White, D.: Framing of project critical success factors by a systems model. *Int. J. Proj. Manag.* **24**, 53–65 (2006)
7. Yeoh, W., Popovič, A.: Extending the understanding of critical success factors for implementing business intelligence systems. *J. Assoc. Inf. Sci. Technol.* **67**, 134–147 (2015)
8. Aziendale, E., Caserio, C., Caserio, C.: Integrated information systems and information systems quality: prospects for analysis and emerging trends. *Econ. Aziend. Online* **10**, 293–320 (2019)

9. Magaireah, A.I., Sulaiman, H., Ali, N.: Theoretical framework of critical success factors (CSFs) for Business Intelligence (BI) System. In: ICIT 2017—8th International Conference on Information Technology, Proceedings, pp. 455–463 (2017). <https://doi.org/10.1109/ICITECH.2017.8080042>
10. Evers, J.M.: Critical success factors of business intelligence and big data analysis. Master Thesis, Tilburg University (2014)
11. Olszak, C.M., Ziemia, E.: Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper Silesia, Poland. *Interdiscip. J. Inf. Knowl. Manag.* **7**, 129–150 (2012)
12. DeLone, W.H., McLean, E.R.: Information systems success: the quest for the dependent variable. *Inf. Syst. Res.* **3**, 60–95 (1992)
13. Floridi, L.: Information quality. *Philos. Technol.* **26**, 1–6 (2013)
14. Yeoh, W., Koronios, A.: Critical success factors for business intelligence systems. *J. Comput. Inf. Syst.* **50**, 23–32 (2010)
15. 4E, Esg 2015. Standards and Guidelines for Quality Assurance in the European Higher Education Area (2015). https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
16. A3ES. Auditoria dos Sistemas Internos de Garantia da Qualidade nas Instituições de Ensino Superior: Manual para o processo de auditoria (2011)
17. Bichsel, J.: Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations (2012)
18. Kabakchieva, D.: Business intelligence systems for analyzing university students data. *Cybern. Inf. Technol.* **15**, 104–115 (2015)
19. Javier, K.: Higher education meets business intelligence. Student Theses, Pap. Proj. (Computer Sci.) **1**, 3–22 (2017)
20. Ong, V.K.: Business intelligence and big data analytics for higher education: cases from UK higher education institutions. *Inf. Eng. Express* **2**, 65–75 (2016)
21. Villegas-Ch, W., Palacios-Pacheco, X., Luján-Mora, S.: A business intelligence framework for analyzing educational data. *Sustainability* **12**, 1–21 (2020)
22. Jayakody, J., Perera, I.: Enhancing competencies of less-able students to achieve learning outcomes: learner aware tool support through Business intelligence. In: Proceedings 2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE 2016), pp. 154–160 (2017). <https://doi.org/10.1109/TALE.2016.7851787>
23. Qlick Sense Homepage. <https://www.qlik.com/us/products/qlik-sense>
24. Porter, M.: Competitive Strategy: Techniques for Analyzing Industries and Competitors. The Free Press (1980)

An Application of Open Data in Public Administrations: The Lanzarote Tourism Indicator Dashboard



Christian González-Martel and José Manuel Cazorla-Artilles

Abstract Increasingly public administrations provide their data under the open data umbrella. It is necessary to develop tools that take advantage of the full potential of new information resources. In this work, we developed a package for R that provides a collection of functions to retrieve, download and manipulate the dataset available by the Canary Institute of Statistics (Instituto Canario de Estadística, ISTAC) through the ISTAC BASE API. In addition, a Shiny web application was designed for the improvement of tourism information to Lanzarote. The collected data in this Shiny web application is related to demand and supply of tourism in Canary Islands and especially in Lanzarote.

Keywords Economic · Databases · Dashboard · Shiny · Tourism

1 Introduction

Open Knowledge [1] define open data as “data that can be freely used, reused and redistributed by anyone—subject only, at most, to the requirement to attribute and sharealike”. During the last years, public administrations have been concerned with offering citizens their data in the open.

Several works have studied how the public consortium deal with open data.

Ruijter et al. [2] conclude that a context-sensitive open data design helps the transformation of raw data into meaningful information. Thorsby et al. [3] research on features and content of open data portals in American cities. Their results show that, in general, the portals were in an initial stage of development and need to

C. González-Martel (✉)

Department of Quantitative Methods, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain

e-mail: christian.gonzalez@ulpgc.es

J. M. Cazorla-Artilles

Department of Applied Economic Analysis, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain

Research Group in the Economy of Infrastructures, Transport and Tourism, Las Palmas, Spain

improve user help and analysis features as well as inclusion of features to help citizens understand the data, such as more charting and analysis.

On the other hand, and linked to open data, there is a tendency, in the academic world, for research to be reproducible. Reproducibility requires that researchers make code and data available so that the data can be analysed in a similar manner as in the original publication. It implies that code must be available to be distributed, data must be accessible in a readable format, and a platform must be available for widely distributing the data and code. In addition, both data and code need to be licensed permissively enough [4]. The complete analytical workflows, fully replicable and transparent, that span from raw data to final publications can benefit from the availability of algorithmic tools to access and analyse open data collections [5, 6]. Dimou et al. [7] present a case of the use of publishing research metadata as linked open data and creating interactive visualizations to support users in analysing data in a research context.

However, due to the data provided in open access are not in a standardized format, it should be adapted in order to accommodate variations in raw data formats. The aim is that end-users can avoid repetitive programming tasks and save time allowing simplification, standardization and automation of analysis workflows facilitating reproducibility, code sharing and efficient data analytics. Following this idea, within the ecosystem of R, several packages have been created to work with data from Food and Agricultural Organization (FAO) of the United Nations (FAOSTAT [8]), World Bank (WDI [9], wbstats [10]), Open Street Map (osmar [11]) among others.

The ISTAC is the central organ of the autonomous statistical system and official research centre of the Government of the Canary Islands. Among its functions are to provide statistical information and coordinate the public statistical activity of Canary Island autonomous region. It provides a rich collection of data, including thousands of datasets on Canarian demography, health, employment and tourism and other topics in an open data format.

ISTAC has made an effort to make the data accessible. The main one is through a web-based graphical user interface (GUI) from where the data can be consulted and downloaded in alternative formats. This access method is tedious for large selections and when the user must access to data very frequently. Other method uses an application programming interface (API) that can be embedded in a computer code to programmatically extract data from ISTAC. In this work, a R package has been developed. It integrates the API into the code and allows downloaded data to be directly manipulated in R.

Based on this package, we have also created a Shiny application that allows a visualization of ISTAC data related to tourism indicators in Lanzarote. In this application, the way to present the information is mainly through graphics. The visualization characteristics are one of the most important features in analysing information from open data sources. Chen and Jin [12] have proposed a data model and application procedure that can be applied for visualization evaluation and data analysis in human factors and ergonomics. Jones et al. [13] research innovative data visualization and sharing mechanisms in the study of social science survey data on environmental issues in order to allow the participatory deliberation. Kao et al. [14] show how to

use a visualization analysis tool for open data with the aim to verify whether there exists sensitive information leakage problem in the target datasets.

This paper provides an overview of the core functionality in the current release version. A comprehensive documentation and source code are available via the package homepage in Github (<https://github.com/rOpenSpain/istacbaser>). The package is part of rOpenSpain, an initiative whose objective is to create R packages to exploit open data available in Spain for reproducible research.

This paper is structured as follows: firstly, we explain the data extraction procedure implemented in the R library and the workflow to achieve visualization of data. In Sect. 3, we explain the architecture of the Lanzarote Tourism Indicator Dashboard. Finally, we present some concluding remarks.

2 The Extraction Routine in Istacbaser

To install and load the last release version of istacbaser, the user should type in R the installation from GitHub command from the remotes package.

```
remotes::install_github("rOpenSpain/istacbaser")
library("istacbaser")
```

When the package is loaded the metadata of each dataset available by ISTAC BASE API are also loaded into the cache variable. It contains information about the title, topic, subtopic, the url to access to the json data, among other. For searching about a specific term, the istacbase search function is provided.

```
busqueda.paro <- istacbase_search("paro",
                                fields = "datos publicadosI")}
```

This seeks among all the ISTAC BASE datasets those in which the pattern “paro” appears within the field “datos publicados I”. Other fields can be “titulo” (default), “tema”, “subtemaI”, “subtemaII”, “datos publicados I”, “origen” and “encuesta”. You can obtain the list of fields with names(cache).

The pattern can be used with regular expression operators. The output is the rows or row of cache that keep to the pattern. Values in the ID column of the output provide data identifiers for subsequent download commands. For example, we can filter the ID of the dataset that collects the registered unemployment according to sex and age groups by islands and months using

```

selectedID <- busqueda.paro %>%
  select(titulo,ID) %>%
  filter(grepl("sexo",titulo) &
         grepl("edad",titulo) &
         grepl("Islas",titulo)) %>%
  pull(ID)

## [1] "emp.est.na.ser.4866"

```

2.1 Downloading Data from ISTAC

We retrieve the data from the dataset with the ID reference using the ISTAC BASE API.

```
df <- istacbase(selectedID)
```

By default, the function `istacbaser` works with human-readable labels. With the argument `label = FALSE` the function transforms the labels into less readable codes.

The indicators in the ISTAC open data service are typically available as annual time series grouped by islands, but sometimes at a different granularity or geographic levels. If the dataset has the “Islas” column, it can be filtered by islands using the argument `islas = TRUE`; otherwise, this argument is ignored. Valid values for islands are: El Hierro, La Palma, La Gomera, Tenerife, Gran Canaria, Fuerteventura and Lanzarote.

The function allows filtering the dataset by dates using the arguments `startdate`, `enddate` and `mr.v`. The argument `freq` controls the granularity of the data for fetching yearly (“anual”), biannual (“semestral”), quarterly (“trimestral”), monthly (“mensual”), bi-weekly (“quincenal”), weekly (“semanal”) values.

2.2 Data Visualization

`Istacbaser` by itself does not have a dedicated function to plot the data but you can use the potential that R provides to visualize the data retrieved. Figure 1 shows the result of the combination of `istacbaser` and `ggplot2` package [15].

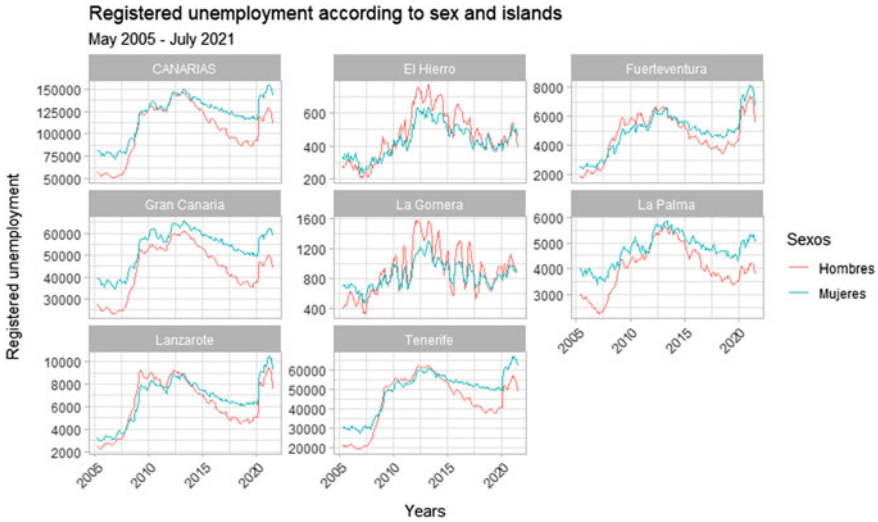


Fig. 1 Data obtained through istacbaser. Visualization with ggplot function from ggplot2 package. Source ISTAC (2018)

```
ggplot (df %>% filter('Grupos de edad' == "TOTAL" &
                    Sexos != "AMBOS SEXOS"),
        aes(x = fecha, y = valor, colour = Sexos)) +
  geom_line() +
  facet_wrap(~Islas, scales="free_y") +
  theme_light()+
  theme(axis.text,x = element_text(angle = 45, hjust = 1)) +
  labs(x = "Years",
       y = "Registered unemployment",
       title = "Registered unemployment
              according to sex and islands",
```

Because that most of ISTAC dataset contain geographical information, map visualization can be represented in a very natural way.

3 Dashboard for Lanzarote Tourism Data Visualization

To provide tourism information to the Lanzarote stakeholders, the Centro de datos de Lanzarote, which is the department of the Cabildo de Lanzarote, which provides with statistical information to Lanzarote and the Universidad de Las Palmas de Gran Canaria establish a collaboration to develop a Shiny web application where the relevant tourism indicators are shown in a friendly way to the stakeholders.

The dashboard can be accessed online.

(<https://datosdelanzarote.shinyapps.io/turismodashboard/>).

The dashboard is based on the idea from the New Zealand Tourism Dashboard. In that sense, the Lanzarote Tourism Indicator Dashboard was developed. The code is written in R statistical software (R Foundation for Statistical Computing, Vienna, Austria) with special attention to the Shiny [16] package, which it is used as a framework for javascript. The Web resource is hosted in shinyapps servers. The maintenance is carried out by Centro de Datos de Lanzarote. The Lanzarote Tourism Indicator Dashboard allows the user to download data and export visualizations. The main visualizations of this Shiny web application are made with the dygraphs package [17], the billboard package [18] and leaflet package [19].

The package *istacbaser* was essential in the development of this application because it is the connector between ISTAC data and the Lanzarote Tourism Indicator Dashboard. All the data it is managed with this package. The *istacbaser* package gets the information that previously ISTAC has obtained by surveys. The Lanzarote Tourism Indicator Dashboard exploits statistics which come from three surveys carried on by ISTAC: Tourism Expenditure Survey, Tourist Movement on Borders Survey (FRONTUR) and Tourist Accommodation Occupancy Survey.

The Shiny web application must keep informed the tourism stakeholders about the sector situation. The information provided by the Lanzarote Tourism Indicator Dashboard is about tourism demand and supply in Lanzarote.

The tourism demand keeps the focus on data about arrivals, expenditure, travel characteristics and tourist profile. The arrivals evolution from the main origin markets it is shown by different time series graphs, and the information comes mainly from the Tourist Movement on Borders Survey (FRONTUR). The evolution could be observed according to a monthly or annual periodicity. The user could also see the values in level, as monthly average of 12 months or as annual variation. There are several options to analyse the results (see Fig. 2, as an illustration), according to the number of nights, the origin market or kind of accommodation.

The expenditure information put the focus on the total expenditure, the expenditure by tourist and the expenditure by tourist and day. The expenditure evolution from the main origin markets comes from the Tourism Expenditure Survey. The evolution could be analysed with annual or quarterly frequency. There are several filters to analyse the origin market even with a breakdown by item of expenditure.

The travel characteristics information is focused on the valuation of aspects to choose the destination, the type of accommodation, length of stay and activities made by the tourists. The travel characteristics evolution from the main origin markets comes from the Tourism Expenditure Survey. The evolution could be analysed with annual or quarterly frequency.

The tourist profile is also analysed to describe tourist characteristics in terms of gender, age, occupation and income. The tourist characteristics information comes from the Tourism Expenditure Survey. The evolution of the tourists according to these characteristics could be analysed with time series graphs with annual or quarterly frequency. Also, a customer journey map was developed to describe by period and origin market the tourist characteristics and the behaviour before and during the travel and the valuation after the travel (see Fig. 3, as an illustration).



Fig. 2 Screenshot of “Llegadas” tab in the “Panel de Indicadores de Turismo de Lanzarote” web resource. Accessed: 2021-10-09

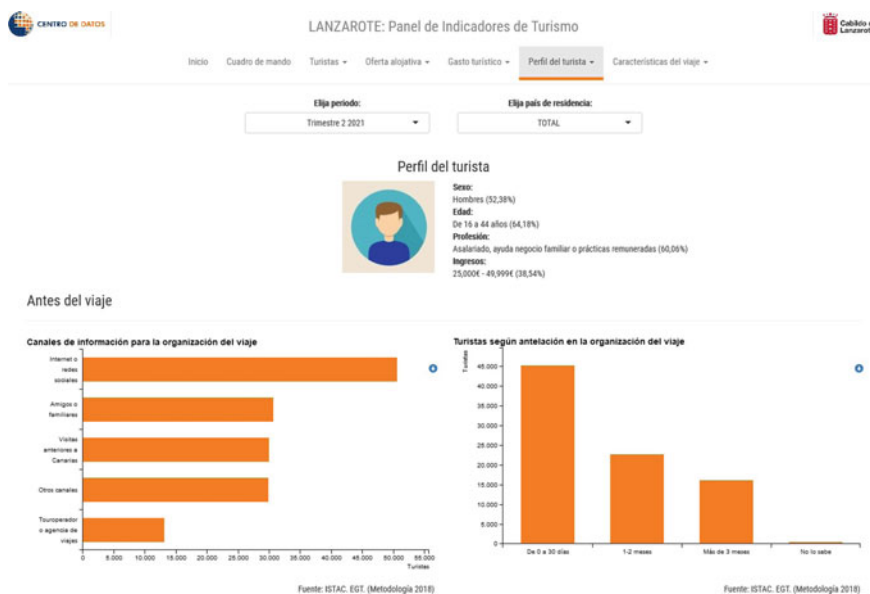


Fig. 3 Screenshot of the “Perfil del turista” tab in “Panel de Indicadores de Turismo de Lanzarote” web resource. Accessed: 2021-10-09

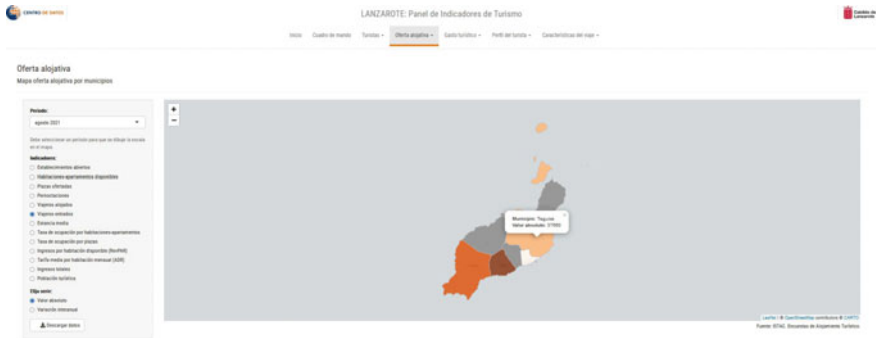


Fig. 4 Screenshot of the “Oferta alojativa” tab in “Panel de Indicadores de Turismo de Lanzarote” web resource. Accessed: 2021-10-09

On the supply side, the focus is on occupation, accommodation capacity, revenue and tourism employment. The supply information comes from the Tourist Accommodation Occupancy Survey. The evolution of tourism supply can be analysed with annual or monthly frequency with observations in level or annual variation for different regions in Lanzarote, according to the type of accommodation or the origin market. The supply side of this Shiny web application shows indicators, such as number of establishments, number of rooms, occupancy rates, average daily rate (ADR), total revenues and touristic employment among others (see Fig. 4, as an illustration).

Within the “Cuadro de mando” section, the user can see at a glance the situation of the tourism sector in Lanzarote by period. The indicators of tourism supply and demand are summed up in just one tab. In concrete, the indicators shown are arrivals, expenditure by tourist and day, length of stay, occupancy rate, number of nights and ADR. Also, there are graphs to show the breakdown of arrivals by origin market and time series graphs for the evolution of arrivals, ADR and expenditure by tourist and day.

4 Conclusions

Nowadays, the public sector is becoming aware of the need their data should to be open. This is useful for researchers and experts in different areas but the heterogeneity of data formats, their complexity and continual updates forces the use of tools that require computational and data visualization expertise to extract and display information. ISTAC offers their statistics under open API and web services.

The objectives of this project were mainly two. First, developing a R package, *istacbaser*, which manages the databases accessible by the API ISTAC BASE. This package allows handling the data provided by ISTAC programmatically for direct use or to include it in other developments.

Second, develop a web service to visualize in real-time the data given by ISTAC. Cabildo de Lanzarote through Centro de Datos de Lanzarote have financed a project where this vision is applied. The result is the Lanzarote Tourism Indicator Dashboard. At this point, the web applications are and improvement to the static reports. A web application lets the user to interact and obtain different responses according to the configuration established. This is quite useful in case of non-technical users because they do not need knowledge about programming or data management, they can obtain insights just by the use of the configuration of the application. This service provides to professionals in the sector all the information related to tourism gathered in one place for consultation and analysis, thereby allowing the public institution to achieve one of the main objectives of the public administration (Fig. 5).

As the reader can see, these projects of dissemination of information can be possible only when the open data is available and as much data is available more information can be offer to the society. As limitations, the procedures for reporting tourism data need to be standardized. Current practices for aggregating data generally involve combining several surveys meaning different granularity. Despite the differences in each source, this can be managed informing, in the application, the source and granularity in order the user knows which source are using at the moment of the query.

In conclusion, we developed the istacbaser package and the Lanzarote Tourism Indicator Dashboard to communicate up-to-date tourism information to the professional and general public. The web application’s pipeline was developed to be extendable.

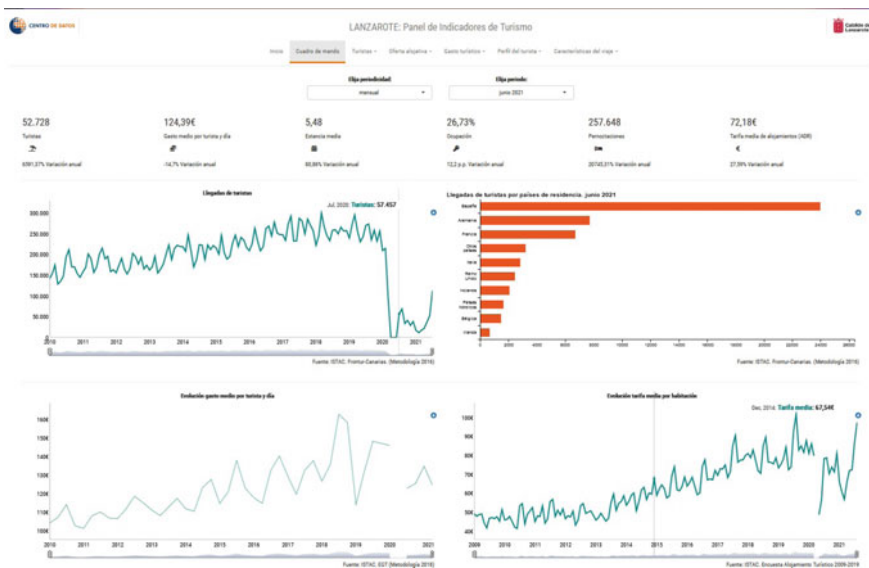


Fig. 5 Screenshot of the “cuadro de mando” tab in “Panel de Indicadores de Turismo de Lanzarote” web resource. Accessed: 2021-10-09

References

1. Open Knowledge (ed.): The open data handbook. Open Knowledge (2015). <http://opendatahandbook.org/guide/en/>
2. Ruijter, E., Grimmelikhuijsen, S., Meijer, A.: Open data for democracy: developing a theoretical framework for open data use. *Government Inf. Q.* **34**(1), 45–52 (2017)
3. Thorsby, J., Stowers, G.N., Wolslegel, K., Tumbuan, E.: Understanding the content and features of open data portals in American cities. *Government Inf. Q.* **34**(1), 53–61 (2017)
4. Stodden, V., Leisch, F. (eds.): *Implementing Reproducible Research*. Taylor & Francis, Andover, England, UK (2018). <https://www.taylorfrancis.com/books/edit/10.1201/9781315373461/implementing-reproducible-research-victoria-stodden-friedrich-leisch-roger-peng>
5. Gandrud, C.: *Reproducible research with R and RStudio*. Chapman and Hall/CRC (2018)
6. Boettiger, C., Chamberlain, S., Hart, E., Ram, K.: Building software, building community: lessons from the rOpenSci Project. *J. Open Res. Softw.* **3**(1) (2015). <http://openresearchsoftware.metajnl.com/articles/10.5334/jors.bu/> <http://openresearchsoftware.metajnl.com/articles/10.5334/jors.bu/print/>
7. Dimou, A., De Vocht, L., Van Grootel, G., Van Campe, L., Latour, J., Mannens, E., Van De Walle, R.: Visualizing the information of a linked open data enabled research information system. *Procedia Comput. Sci.* **33**, 245–252 (2014)
8. Kao, M.C.J., Gesmann, M., Gheri, F.: FAOSTAT: download data from the FAO-STAT database (2020). <https://CRAN.R-project.org/package=FAOSTAT>. R package version 2.2.1
9. Arel-Bundock, V.: WDI: world development indicators and other World Bank Data (2021). <https://CRAN.R-project.org/package=WDI>. R package version 2.7.4
10. Piburn, J.: *wbstats: Programmatic Access to the World Bank API*. Oak Ridge National Laboratory, Oak Ridge, Tennessee (2020). <https://doi.org/10.11578/dc.20171025.1827>
11. Eugster, M.J.A., Schlesinger, T.: osmar: OpenStreetMap and R. *R J.* **5**(1), 53 (2012). <http://osmar.r-forge.r-project.org/RJpreprint.pdf>
12. Chen, X., Jin, R.: Statistical modeling for visualization evaluation through data fusion. *Appl. Ergon.* **65**, 551–561 (2017). <https://pubmed.ncbi.nlm.nih.gov/28110916/>
13. Jones, A.S., Horsburgh, J.S., Jackson-Smith, D., Ramirez, M., Flint, C.G., Caraballo, J.: A web-based, interactive visualization tool for social environmental survey data. *Environ. Modell. Softw.* **84**, 412–426 (2016)
14. Kao, C.H., Hsieh, C.H., Chu, Y.F., Kuang, Y.T.: Using data visualization technique to detect sensitive information re-identification problem of real Open Dataset. In: *Proceedings—2016 International Computer Symposium, ICS 2016*, pp. 330–334 (2017)
15. Wickham, H.: *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York (2016). <https://ggplot2.tidyverse.org>
16. Chang, W., Cheng, J., Allaire, J., Xie, Y., McPherson, J.: *shiny: Web Application Framework for R* (2020). <https://CRAN.R-project.org/package=shiny>. R package version 1.5.0
17. Vanderkam, D., Allaire, J., Owen, J., Gromer, D., Thieurmel, B.: *dygraphs: Interface to 'Dygraphs' Interactive Time Series Charting Library* (2018). <https://CRAN.R-project.org/package=dygraphs>. R package version 1.1.1.6
18. Perrier, V., Meyer, F.: *billboarder: Create Interactive Chart with the JavaScript 'Billboard' Library* (2020). <https://CRAN.R-project.org/package=billboarder>. R package version 0.2.8
19. Cheng, J., Karambelkar, B., Xie, Y.: *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library* (2019). <https://CRAN.R-project.org/package=leaflet>. R package version 2.0.3

Structuring Best Practices of Search Engine Optimization for Webpages



Riaze Miguel Ribeiro Issá  and Jose Paulo Marques dos Santos 

Abstract The main objective of search engine optimization (SEO) is to present the most relevant information to the user, so that his necessities may be satisfied. Therefore, it becomes imperative to understand and be familiarized with the constant developing algorithms, on what mainly concerns the variables it gives the most importance to. In this article, an in-depth explanation is presented, from a basic to a more advanced level, of the internal and external variables required for a good positioning of website pages, enabling them to escalate in the search engine results page (SERP). Such holistic approach is useful, so that the reader, using the most appropriate techniques and tools, may be one step ahead of the remaining competition.

Keywords Search Engine Marketing (SEM) · Search Engine Optimization (SEO) · Search Engine Results Page (SERP)

1 Introduction

In a digital environment that is under constant development and rapid expansion, where just one Internet connection provides unlimited access and the sharing of massive amounts of information, where there are millions of websites in which the daily increment of information is immeasurable, increasing difficulty in finding reliable and quality information, the need of creating relevant pages for user queries is of pivotal importance. The strive for a digital presence that is visible to the user has been driven by search engines that feed off this struggle which, in turn, fosters their development. The search engine algorithm will always give priority to the pages of optimized websites and, therefore, the strive for top and prominent rank positioning is achieved by websites with the best optimized pages for certain queries.

R. M. R. Issá · J. P. Marques dos Santos
Universidade Europeia, Lisboa, Portugal

J. P. Marques dos Santos (✉)
University of Maia - ISMAI, Maia, Portugal
e-mail: jpsantos@umaia.pt

This paper presents recommendable search engine optimization practices and shows that, when used properly, they make it possible for a website to achieve a relevant position in relation to its competitors. The concepts presented in this study stem from the basic to the more advanced by means of a more comprehensive approach to SEO, by defining what it is and once the basics are covered, it provides details and in-depth information on the potential strategies that can be applied inside or outside websites, aiming at improving the visibility of pages while maintaining consistency in their optimization.

2 Search Engine Marketing—SEM

The main objective of search engine marketing is the development of strategies that increase traffic on certain website pages, whether these are organic—SEO, or by means of ads—SEA [3]. SEM can be highly complex due to the large number of factors that influence its behaviour, whether these are the user or the search engine algorithm. The volume of searches for a given page has a very significant impact on the moment of positioning [7].

The organic search results organized by the search engine results page (SERP) are ordered by their relevance according to the user's choice of keywords. It is important to know how to distinguish organic results from paid results. Paid results are strategies developed in which the marketer pays the search engines to be displayed in ad format at the top of the page. The goal of any user is to find relevant information according to his query that can be broken into different categories: navigational, transactional or informational. The most important fact when creating a SEO digital marketing strategy for a website is to develop it with a deep understanding of the target audience. Understanding how most user searches and target market searches work can help to have a more effective impact on the user and, subsequently, to find ways to target new leads [13].

Query formulation can be carried out by using words or short phrases. Most users formulate their queries by using one to three words. When the search engine returns the results of a query on the SERP, it is possible to determine if the contents are the desired ones by observing these results or in other words, if they correspond to the expected response to the query or if on the contrary, they require users to refine their search until the desired results are obtained. Thus, when users find the content they need, this query translates into a satisfying experience because the user was able to complete the intended task [13].

3 Search Engine Optimization—SEO

Search engine optimization is, by definition, one of the most effective marketing strategies in gaining organic traffic by means of processes such as writing, programming and design. All these efforts help in the gradual increase of the volume, quality and visibility of any website page which users can access by means of a search engine [22]. Williams [43] states that SEO, search engine optimization, includes all the steps that a webmaster (whoever designs and maintains a website) takes to ensure the visibility of their pages on Google's results page, SERP. SEO strategies help in the visibility of digital platforms not only from a technical perspective, but also in an equally creative way that helps improve the ranking position in organic searches and achieve high rankings on SERP's [24].

The creation of strategies and their implementation in SEO are not static processes. While observing results, the webmaster must resort to his know-how and flexibility to change or update optimization techniques. This constant transformation which takes place by means of experimentation must meet previously established objectives. As such, this manipulation of a website's characteristics should be considered to improve its positioning on the SERP in relation to competitor pages [25]. Within this scope, Zilincan [47] states that SEO is characterized as the set of modifications and techniques that help search engines to collect, index and interpret the contents of a website. Zhang and Cabage [45] explain that within search engine optimization, SEO is characterized as a set of techniques, namely content marketing, link building and social sharing. A well-developed SEO strategy is critical to the success of online businesses.

According to Chotikitpat et al. [9], the search engine algorithm rules must be strictly followed in order for page contents to be properly indexed. With this traffic-boosting art inside a website, there will be new visitors and leads and, in turn, not only will increase user engagement but also brand perception. Once users see the reputation within the brand or, in other words, the authority within the segment, this will boost the probability of an increase in the volume of sales.

4 Google's Guidelines

In this SEO reality, not all strategies and techniques are consensual, and the scientific community is divided between what is to be considered good practices and what is not. There are webmasters who follow Google's guidelines and webmasters who use algorithm manipulation to achieve their goals.

White hat SEO include all strategies that aim to place a page in a high-ranking position by always following the guidelines, i.e. the SEO strategies approved by Google itself [43]. These are techniques that contribute to the ranking of a page [23] in the form of quality content, engagement and user-friendly navigation. According to Berman and Katona [6], White hat SEO contributes to the content of a website

effectively and positively which, in turn, translates into consumer satisfaction and consequently, makes the website more relevant from the user's and the Google robot's point of view.

On the other hand, black hat SEO encompasses all the manipulation and exploitation of Google loopholes to position pages with the risk of being penalized or banned if these practices are discovered [43]. Accordingly, Killoran [23] characterizes this type of SEO by the set of unconventional initiatives that boost traffic, namely the use of spam and irrelevant or thin content. Moreover, Aswani et al. [3] add that this way of earning more traffic has its dark side considering that in the long run, black hat SEO is ephemeral due to the constant regulation and updating of Google's algorithm mechanisms, which identify and penalize these types of strategies. According to Zhang and Cabage [45], these types of penalties result in the demotion of the website on SERP positions and result in fewer impressions, or the website may also be banned and blacklisted by Google in a worst-case scenario. Most websites that are caught by the search engine have a kind of black hat, which can include practices of keyword stuffing, duplicate content, spamming, link farming, cookie stuffing and doorway pages [20].

The grey hat SEO lies between these two extremes. Although not approved by Google, these types of SEO strategies have, unlike black hat SEO, a low probability of being penalized. Grey hat tactics have a higher risk than white hat tactics, but not as high as black hat tactics [43].

5 Long Tail Keywords and Keyword Research

Anderson [2], the creator of the long tail concept, states that there is a clear division between two product realities: mass market products and niche market products.

According to Enge et al. [13], there are search terms that receive more than five thousand searches per day but correspond to only 30% of all searches carried out on the Web—long tail searches, as opposed to the most common searches, short tail searches that are 70% of all Web traffic. The keywords long tail add more relevance to a given user's query because they are the most descriptive and precise and therefore confer more value to the traffic generated for the pages in question. Understanding the demand for each search is pivotal, considering that there is a need to create long tail content, i.e. niche content that is not searched for by all users, but by more specific users who are interested in the topic. Before starting any SEO project, it is necessary to establish the keywords target because the entire content and optimization strategy will stem from them. Market trends may be understood through analyses, as well as potential trends within the segment, because they make it possible to foresee the increase in demand for certain products. This way, it will be easier to find the best search terms and meet consumers' search needs. Considering that any search term that is typed into the search engine is recorded, the use of keyword search tools is vital. Although important, these tools do not directly show how valuable or important a keyword really is, as it is necessary to understand the two variables of a keyword,

its difficulty (keyword difficulty—KD) and its search volume (SV). Also, important to note is the metric scale of these two variables which in the case of keyword difficulty, goes from 0 to 100%, and in the case of search volume, which in theory can be unlimited, can start at 0 and exceed 100%.

When selecting keywords, there are aspects to take into account. A search that is carried out very randomly or broadly may cannibalize other keywords. This process is counterintuitive because it does not show an internal hierarchy on the pages. Consequently, the search engine crawler is forced to make its choice between what is more and less relevant, creating a serious optimization error. Therefore, search, hypothesis formulation and testing must be targeted at the SEO goal [13]. Within this scope, Ramlall et al. [35] state that a short tail strategy is more challenging than a long tail strategy, considering that short tail keywords have more competition due to fact that they facilitate queries as opposed to long tail keywords, which are more specific, have less search volume and queries are less common. This way, an ecosystem of billions of unique queries is built [24]. Webmasters should always carry out keyword analyses from the users' point of view so that there is a greater degree of understanding of the keywords and, in turn, insert them into the SEO strategy, thus ensuring qualified traffic to the website pages [41].

6 On-Page Optimization

As described by Yalçın and Köse [44], on-page optimization consists of all website improvement tactics that are carried out internally. The goal is to make sure that the entire page construction of a given website is readable to search engine indexers or in other words, the capture and transformation of HTML and CSS code into data that describe and rank this Web property to perfection. Thus, Ghulam et al. [14] state that there are a set of variables which together make up the on-page optimization of any website. As such, a correct and complete on-page optimization depends on tree structure, the selection of the name for the domain, the page name, the title metadata, the site description, the meta keywords, the organization of the URLs, the heading tags, image alt texts, the written content on the pages, website navigation, its sitemap and robots.txt status.

In the hierarchical structure of a website, pages should be organized in a top-down format with a static URL strategy that facilitates their indexing in search engines as for example, the search carried out by users. There must also be a semantic relationship with the brand and the name of the product or service for which it is targeted, hence the use of hyphens or addition symbols to separate the terms identifying the subpages.

6.1 *Title Tags*

The function of a page title or title tag is that of being merely indicative of the themes and sub-themes of the pages, as well as to help guide users in their search. The title tag is ranked by the search engine algorithm, so it is critical that the theme of the page correlates with the user's search context (UI) [33]. According to Enge et al. [13], to achieve this optimization, the length of the title tag must be limited and no shorter than 50 to 60 characters long. There should be no duplicate title tags and the main keyword must always be present, ideally as far to the left as possible [46].

6.2 *Meta descriptions*

According to Khraim [22], meta descriptions are visible to the eyes of a consumer but they should focus on the search engine and its respective indexing. By norm, the function of meta descriptions is to provide a summary of a page's contents in order to persuade the user to click on the page.

According to Enge et al. [13], considering that a meta description is created by the webmaster, their writing should be clear, objective and promotional within a minimum of 70 characters and a maximum limit of 155 characters to be able to boost website traffic. The correlation between the user's query and the keywords inserted in a meta description will not only allow the search engine to read their content, but also to rank the most relevant pages according to the UI for the user within its indexing process [46].

According to Williams [43], this variable, as well as the page title, is primarily responsible for increasing click-through-rate (CTR). The more clicks the page receives, the better its ranking position on the SERP and more importantly, the satisfaction of a user's search need is achieved. A page that is in the first position receives, on average, about 31% of all clicks generated on that page on the SERP. Next, the page in the second position receives on average 14% of all clicks, while the page in third position receives 10% of clicks. On the other hand, pages on the second page of the SERP receive about 4% of clicks and those on the third page receive only 1.6% of clicks.

It is important to understand the importance of good ranking position. The higher it is the better, considering that a drop in one or two positions is enough for a significant decrease in traffic. For example, pages between the sixth and tenth position on the first page of the SERP have a CTR of only 3.73%, which is less than the CTR of the page in fifth position which is about 5%. [43].

6.3 *Meta keywords*

When Google's algorithm first appeared, webmasters discovered that the search engine considered meta keywords to be very relevant, and they were one of the key ranking factors. Thus, they began to over-optimize this field by means of keyword stuffing practices to manipulate rankings on the SERP. However, Google identified this practice and started ignoring this tag although not completely because it is still used to identify spammers. It is currently considered a good practice to leave the meta keywords field empty [43].

6.4 *Heading Tags*

Heading tags are titles, and their function is to organize and rank the content within a page. These titles and subtitles must reflect the content of the page to inform the algorithm and grab the user's attention, because the more interesting the content, the longer the user will stay on the website. As stated by Busche [8], Heading tags serve to provide greater organization and sequence of titles and subtitles, both for the user and for the Google crawler. The H1 or Heading tag 1 is the main and most relevant title, so there can be no other H1. It is unique. The subheadings that follow range from H2, H3, H4, H5 and H6, the latter being the least important on the hierarchical scale. Heading tags play an essential role in organizing content delivered to the user [13]. Heading tags should contain the main keyword, mainly the H1, H2 and H3 as they are the second most important on-page factor [46].

6.5 *Images and Media Alt-Text*

According to Bailyn [4], search engines are only able to read text. During their indexing process, when the crawler identifies an image or other type of media file such as videos, this type of content is discarded.

With major technological advances leading to the consumption of other types of content other than text, alternative texts (alt-text) were created. They are basically tags that are used to describe non-text content to the crawler. For proper optimization, the alt-text should contain the keyword or descriptive phrases of the page topic [46]. In agreement, Enge et al. [13] stress that search engines only accept HTML language and refuse any other type of format present in websites. However, by placing Alt attributes, it becomes possible to name attributes such as images, audios, videos, Adobe Shockwave files or even another type of computer language such as Java, so that the algorithm can understand the content exposed in the digital property. Choudhari and Bhalla [10] indicate there is a specific way to optimize videos called video search engine optimization (VESO). This optimization takes variables such

as the title, the description, the transcribed script, the thumbnail which describes the content with an image, as well as annotations that promote linkbuilding through external links. As such, there must be thorough knowledge of the market segment and the target to whom you want to pass the message to, so it is possible to make extraction of keywords more effective because these are the ones that have a higher search volume of the target audience.

6.6 Taxonomy and Ontology

Enge et al. [13] state that taxonomy is essentially a two-dimensional hierarchical model of a website's architecture. It is essential in the presentation and arrangement of the website's themes to the consumer because the better this architecture is designed, invariably, the better the consumer experience will be. According to the same authors, ontology is a way to map the human mind. When people think about a certain topic, it is common to interrelate it with other topics, and this is important to consider when planning the categorization and creation of clusters of similar and correlated topics within the site. In this same line of research, Batsakis et al. [5] defend that it is the taxonomy that allows grouping similar topics or categories to boost the user experience by delivering more content and making it easier for them to navigate on pages.

6.7 Sitemap

The sitemap is the map of a website that contains a cluster of URLs and other files, such as extension files (PDFs, TXTs, DOCs) that make up the virtual property. According to Ziakis et al. [46], the XML file is created by the website's developer, and it is submitted to make it easier for the crawler to find all the subpages of a website. Additionally, this file notifies the search engine when any change occurs at the website level or even certain aspects at the page level. According to Visser and Weideman [42], the XML extension is used to indicate the indexing of pages to the search engine on the website so that they appear on the SERP. This same indication can also reinforce the priority of the pages you want to give more relevance to base on their traffic, considering that the higher the volume of traffic is, the higher its priority on the site should be. However, Enge et al. [13] guarantee that it is not by adding a new URL to the sitemap that its indexing is guaranteed. It is essential for the site to have its sitemap optimized as this allows for the customization of URLs and metadata if they are duplicates or hard for the crawler to collect. In addition, it is also possible to change or choose canonicals.

The sitemap must be present in robots.txt because according to Google [16], the robots.txt file is automatically read by the robots when they arrive at a given website. This file should contain “commands” for the robots, such as pages we want or do not want to index.

6.8 Responsive Navigation, Page Speed and Structured Data

When working on the hierarchical structure of a website, there are two sides to consider, the human’s and that of the robot’s. From the user’s point of view, i.e. the human side, the site must respect how the user usually consumes information, namely the menus in the header, the fact that reading is done from left to right, the webpage layout should avoid too much sensory information for users so that they do not get dispersed or frustrated by not being able to absorb everything on a particular page. Each webpage must be designed with the user in mind, so that they enjoy a good UX, with a simple and optimized navigation and no distractors. From the point of view of the crawler, it must be able to read and assimilate the most important page details. For this to be possible, the tree structure must also be optimized [43].

According to Lali et al. [24], the loading speed of a website’s pages (pagespeed) is a determining factor in the positioning of results on the SERP. Within this scope, Manhas [26] reinforces this line of thought by stating that for a good performance of pagespeed, it is necessary that the internal and external elements of the page in other programming languages—such as HTML, CSS, JavaScript and json—improve aspects related to site navigability. High-performing websites have higher percentages of user retention, engagement and conversions. For Ziakis et al. [46], loading time is an extremely important factor, considering that engines include it in their algorithm. The longer the loading time of a page, the worse its position on the SERP will be.

One element that helps the crawler in its interpretation of page information is structured data. This type of interpretation can be found on the Google results page from a previous HTML construction. This type of classification is a great support in sending information to the search engine so that it can categorize and present the data in the best way possible [24].

6.9 Bounce Rate and Time on Website

The bounce rate happens when the user clicks on one of the SERP results and then returns to the Google results page again. The faster the user returns to the results page, the greater the user’s dissatisfaction is identified by Google, penalizing the reputation of the website. On the opposite side of this experience, we find the user who navigates from page to page on a website increasing their dwell time and informing Google that the webpages are relevant and, therefore, it places them in better positions [43].

In agreement, Ziakis et al. [46] state that websites with low bounce rate inform the search engine that this is a high quality and extremely relevant website for the user.

In addition to Ghulam et al.'s [14] perspective, there are more key variables to take into account when a webmaster analyses the on-page optimization of a given website, namely: the status code of the various URLs that make up the website; the SSL security certificate; the structure of the URLs; checking the status of indexing tags and authority; the depth of the website and the correspondence between the website's language and the programming language.

6.10 Status Code

According to SEMrush [39, 40], HTTP status codes represent all requests sent to a Web server by search engines or platform users. Multiple webpages with 4xx and 5xx codes can negatively affect user and crawler experience, resulting in traffic drops, which in turn lead to loss of organic positions.

When analysing multiple status codes, it is essential to establish a priority in checking the code statuses and fixing existing errors. Priority should first be given to errors that affect the website, be it in the presentation of results to the user or to the search engine, and only then to less harmful codes. As such, the status codes that come up in a status code analysis more frequently are 2xx, 3xx, 4xx and 5xx.

The 2xx status codes indicate that the communication was successful and that the Web server was able to respond to the user's request [31].

The 3xx or redirect codes are an indication for a certain internal URL to redirect to another URL. This can take on seven different names: 301, 302, 303, 304, 305, 306 and 307. However, a special preference should ideally be given to the 301 redirect—permanent redirect, since it is the redirect that ensures that there is absolutely no loss of URL authority. On the other hand, 302 redirects are also quite common. This type of redirect is temporary, indicating to the search engine in question that the user is temporarily being redirected to another page, and thus there is no passing of authority. These should be done with caution, as a slight change will cause discomfort and a bad user experience, which the search engine will detect and take action, according to the severity of the situation [31].

4xx Errors indicate that there was a request to access a certain page or when access to a webpage is restricted. The most common errors are 401—Unauthorized; 403—Forbidden; 408—Request Timeout and 404—Not Found. This type of URL jeopardizes the user experience within the website and penalizes users, considering they cannot access a certain page or file through the link they click on. Additionally, internal links that send users and the search engine crawler to dead end pages imply a spending of the crawl budget (number of pages that are read by the crawler on a daily basis) with a subsequent penalization. Webpages with status code 4xx will be removed from the Google Index [31].

5xx Errors or Server Errors indicate that there are problems with the server. These problems should be immediately reported to the website developer or

hosting provider of the digital platform and resolved with the utmost urgency. The most common errors include misconfigurations, server overload or poor server performance [31].

6.11 SSL Security Certificate

According to Ziakis et al. [46], the acronym SSL stands for secure socket layer, a protocol that establishes a code between the server and the browser which allows information to be transmitted securely. This certificate ensures that the website is trustworthy, that it protects user data and prevents spam tactics. Websites with an installed SSL certificate use the https protocol instead of http. Thus, sites with an SSL certificate tend to rank better on the SERP than others.

6.12 URL Structure

According to Ahrefs [1], in the optimization of URLs, Google's algorithm gives preference to friendly URLs, that is, those with a simplified structure that allows the user to know where the website is located, as well as deliver easy-to-read information to the crawler, thus also improving its loading speed. When a new page is created, a new URL is generated and as such, knowing its structure is vital.

The URL protocol is the first element of analysis, and it is characterized by a colon, two forward slashes, followed by the host name. It can be referred to as a transfer protocol or scheme and will determine how information is transferred between the host and the user who is carrying out the search [1].

The choice of a good domain is essential when creating a website because good accessibility represents the users' gateway to the Internet. During the selection process, the choice for an exact domain strategy (EMD) as Williams [43] states, is a type of strategy that will be penalized by Google itself, due to the unfair advantage positioning it represents. On the other hand, the subdomain is the extension of the primary domain and is used to organize the different content on a website and for redesign functions. It is most often used in blogs and eCommerce stores [1].

The top-level domain (TLD) is a suffix related to the last letters after the period in the URL. There are over 1000 possibilities of TLDs which a webmaster can choose from, considering that the most used are the geographic location ones that are typically related to where the website domain was created [1].

The subfolder and slug are the organization and structuring of the website. By norm, a subfolder is the category pertaining to some broader topic that, in turn, unfolds into more specific content, the slugs. In a blog, a subfolder would be the main page which unfolds into other more detailed webpages on a part of the base topic. In the case of ecommerce, it is the category or parent page, which in turn incorporates other product webpages related to the category described [38].

An idea and good practice in the slug to take into account is to use the keyword of the page and, if there is more than one term, resort to the use of hyphens. This should be short and objective, excluding the use of special characters such as capitals, underlines and numbers [1]. In the same research line, Ziakis et al. [46], state that including the main keyword in the URL is an important positioning factor and that it presents a simplified URL structure, considering that the crawler will identify the webpage theme in a simpler way.

6.13 Noindex and Nofollow Tags

According to Google [15], the content attribute instructs the search engine on how it should crawl and index the pages. If there is no metatag blocking certain webpages in robots.txt, the crawler will interpret the information to index and follow the webpages. By doing this, it is giving itself permission to show the pages on the SERP and read all the links from them. On the other hand, pages with the noindex metatag will never be displayed on the SERP and read by the crawler. However, for the metatag to work, pages that are not intended to be indexed cannot be blocked by robots.txt, otherwise the bot will not read the information [18].

6.14 Crawl Depth

Crawl depth is extremely important for both the user and the crawler. According to SEMrush [39], search engines consider the most relevant pages to be those which, starting at the homepage, require the fewest clicks to be accessed. A good navigation structure allows access to all website pages with a minimum number of clicks. Ideally, no more than three clicks.

6.15 Canonical Tag

The canonical tag provides information to search engines about which version of the page has priority. It must appear in the sitemap and be indexed. Canonicals are extremely useful in situations where the same page is presented with different URLs (e.g., parameters, pagination) or in cases where there are slight content variations (e.g., product variants) [28].

6.16 Hreflang Tag

The hreflang tag is useful when there are multiple versions of pages in different languages, and it indicates to the search engine these different variations and shows the appropriate version of the page according to the language of the given region [37].

The hreflang tag must be in line with the `<html>` lang attribute which, in turn, must be configured on the website [32].

7 Page Rank—PR

There is one essential factor that is a strong indicator in the presentation of results. This technology is called Page Rank, PageRank or simply PR, and it is the algorithm used for Google searches that effectively positions website pages on the SERP. PageRank was named after one of Google's founders, Larry Page. This algorithm is measured on a scale of zero to ten but in fact, no page has rank zero. The page is unrank, which is a value very close to zero. It is important to note that Google can index pages with no rank [43].

When a page links from page A to page B, page B, the destination page increases its PR in proportion to the PR of page A, the source page. If the source page, A has a PR of zero, the destination page, B receives only a tiny portion of its PR. However, if page A has a high PR, page B receives a huge leverage.

If a page has a lot of inlinks, it will receive PR from the pages that are sending those inlinks, making it a more relevant page. A simple way to understand this concept is to consider links as “votes”. The pages with the most votes are the most relevant pages from Google's perspective [43]. The more authority and quality links a page has, the more PR it receives and, in turn, the more relevant it becomes from Google's perspective. In other words, the page gains reputation and status. The more reputation a page has, the better it will rank on Google's SERP. Nowadays, Page Rank is still a topic that is a subject of debate and discussion in the scientific SEO community. Taking into consideration that this is the technology used to position pages, it is essential to understand the metric and the variables that enable this process on Google [43].

Today, the Page Rank information that Google made available when SEO first appeared, namely with its first algorithms, is completely hidden from all users, webmasters included. However, multiple companies have invented tools that make it possible to simulate the search engine algorithm with the aim of coming close to the real authority of the websites.

Among the most successful tools, Moz stands out. Similarly to PR, Moz has a scale that goes from 1 to 100 to evaluate both the authority of a domain (domain authority—DA) and of a given page (page authority—PA). The domain authority—DA is calculated based on several factors, of which we can highlight the number of

links to the domain root and the total number of links [29]. This score is the closest way to reality and is very useful in comparing sites when checking the reputation of a given domain. On the other hand, the DA is, as its name indicates, the authority of each web page. The set of all webpages of a given website is called domain authority, the DA [30].

8 Off-Page Optimization—Link Building

As described by Yalçın and Köse [44], off-page optimization includes all the website improvement tactics carried out externally. The goal of a link building strategy is to connect multiple platforms, namely brand and other platforms, to gain greater awareness and digital maturity, which in technical SEO language is better known as increasing authority. A page with more links, i.e. both internal and external hyperlinks to other websites, is the best signal to the search engine that that page is relevant to users, and therefore has a larger digital footprint, a closer-knit community, a greater link of trust and effectiveness in presenting results [19]. In Killoran's [23] view, working on link building as a way to gain authority involves knowledge of attributes, namely how to gain popularity, relevance and trust for user search. According to Enge et al. [13], the authority of a webpage is the unit of measurement of its reputation. This authority is measured by having multiple links to different websites, resorting to the hyperlink method, where it transfers authority via the different links. As a norm, there should be a relationship between the links, such as interaction sites related to the business area (core business) or social networks, so that this management can be done correctly. A higher authority is correlated with better quality content, since it is this type of content that generates greater social engagement by existing platforms, thus helping the positioning of the page on the SERP. For Deitel [11], the link building strategy is based on three pillars, namely the reciprocal link, where there is a mutual exchange of links; the natural link, in which the link is created through a content marketing strategy (for example, a blog article); and link baiting where there is a fostering of awareness in relation to a certain content.

When comparing link building and social sharing strategies, the link building strategy is undoubtedly the one that generates more positive results, such as greater traffic generation for the website, which results in a growth of authority that, in turn, allows for a better positioning on Google's results page. Social sharing proved to be a less efficient strategy because the volume of traffic was lower when compared to link building strategies [45]. A link building strategy can be approached in different ways (Fig. 1), such as article submission, directories, forum postings, presence on social platforms, blogging (which is based on content marketing), and guest blogging, which is the creation of written content pieces to be published elsewhere with the intention of placing the anchor text and receiving the backlink from the site where it was published [21]. Anchor text is a key element within content pieces which make it possible to link to other pages, either in or out of the website [27].

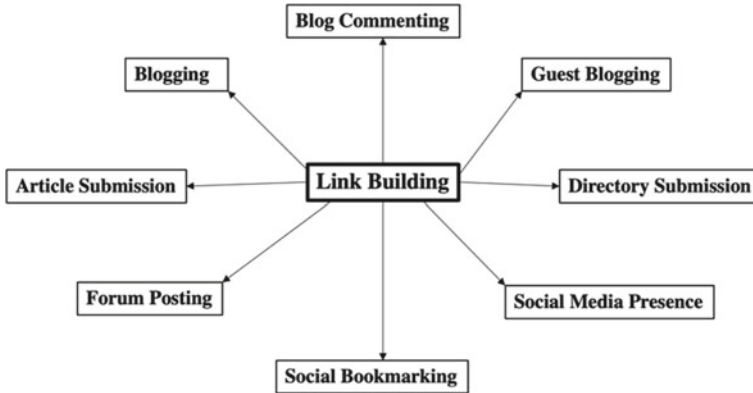


Fig. 1 Link building strategies

For an understanding of domain authority, it is essential to understand how links work. There can be two types of links: internal and external. Additionally, they also have two attributes: nofollow and dofollow. Nofollow links are those which tell the search engine robot (`robots.txt`) not to follow a certain link, thus not passing authority. Unlike nofollow links, dofollow links give the robot permission to follow them, in turn passing authority to the domain [17].

9 Content Marketing

Search engine optimization and content marketing are two inseparable realities that need each other. SEO generates traffic and optimizes content while content marketing creates the content. According to Pažėraitė and Repovienė [34], this form of marketing exists when there is good SEO work supporting its entire structure, considering that the more traffic to a page, the more users will consume the content on that page. In the same line, Rowley [36] states that it is through good pieces of content, that positive emotions are attracted and generated in the user, with the final result being a higher degree of loyalty, satisfaction and trust. In this light, Di Gangi and Wasko [12] state that personalization of communication between users has a significant and positive impact on their social interaction during their navigation, as this engagement is defined as an individualized concern. Moreover, when users receive a type of communication that is personalized and meets their interests, greater satisfaction is inherent, with the final result being that of greater engagement. It is also necessary to consider that interaction quality must be based on the amount of information about the user, and once you have this information broken down and analysed, the greater the likelihood of success in engaging a lead. That is why it is so important to focus on the quality and not the quantity of the message, where in

the second case, the brand may become associated with negative feelings due to its high intrusiveness and lack of interpretation of needs [12].

In SEO, pages that want to rank high in search results need to be unique, authentic and have relevant content. Thus, all quality content must respond to two attributes: it must match search engine queries and it must contain links. Keyword research is essential to ensure that they are searched for by users [46].

The more quality content each page has, the more likely it is to rank correctly in Google. A recommended number would be 500–1000 words per page. Otherwise, when reading the page, the search engine will struggle with missing keywords and have a harder time positioning the pages [37].

When writing content, it is important to understand what the topic of the page is. Next, considering which keywords are best to include in the content, to meet the search carried out by users [41]. A competitor analysis should be carried out, especially those that have a good positioning in search engines, to have an idea of the type of content that generates good results.

At the time of writing, including the keyword of the page with proper sentence coherence is recommended. Keyword density is the number of times a given keyword appears in the text of a page when compared to the rest of the website text and can be calculated by the following formula:

$$\frac{\text{number of times the keyword appears}}{\text{total words}} \times 100$$

It is a key concept to consider when creating content [46].

Another important factor to take into account is the linking of each page of the website. The intention of creating internal linking is not only to make the user's visit more enjoyable, but also to facilitate the paths and indexing of the individual components of the website which consequently leads to improved page positioning [46].

10 Conclusion

We hope that the description of the techniques and tools that aim to improve the positioning of website pages in search engines have contributed to the understanding of the variables that require greater attention, so as to enable the search engine algorithm to carry out a better reading of all the information presented. For such a search, the greater the detail of each optimization, the greater the probability of a better ranking positioning on the SERP in relation to competitors.

The SEO approach must be a holistic one or, in other words, it must take into account several variables in order to understand which ones have the greatest impact and if they are correctly optimized, so as to achieve consistency in the presentation of the results and in turn, in their positioning.

It is essential to monitor results and constantly track page positions because the search engine algorithm is constantly being updated, and it is this monitoring that makes it possible to assess the impact of algorithm updates on all SEO work.

References

1. Ahrefs: How to create SEO-friendly URLs (2021). <https://ahrefs.com/blog/seo-friendly-urls/>
2. Anderson, C.: *The Long Tail*. Hachette Books (2008)
3. Aswani, R., Kar, A.K., Ilavarasan, P.V., Dwivedi, Y.K.: Search engine marketing is not all gold: insights from Twitter and SEO Clerks. *Int. J. Inf. Manage.* **38**(1), 107–116 (2018). <https://doi.org/10.1016/j.ijinfomgt.2017.07.005>
4. Bailyn, E.: *Outsmarting Google*, 1st edn. Que Publishing (2011)
5. Batsakis, S., Petrakis, E.G.M., Milios, E.: Improving the performance of focused web crawlers. *Data Knowl. Eng.* **68**(10), 1001–1013 (2009). <https://doi.org/10.1016/j.datak.2009.04.002>
6. Berman, R., Katona, Z.: The role of search engine optimization in search marketing. *Mark. Sci.* **32**(4), 644–651 (2013). <https://doi.org/10.1287/mksc.2013.0783>
7. Brooks, N.: The Atlas rank report : how search engine rank impacts traffic which factors determine click volume ? Atlas Institute, 1–4 (2004). <http://94.126.173.33/ad2006/adminsc1/app/marketingtecnologico/uploads/Estudos/atlasonepoint-howsearchengineimpactstraffic.pdf>
8. Busche, L.: *Powering Content*. O'Reilly Media, Inc. (2017)
9. Chotikitpat, K., Nilsook, P., Sodsee, S.: Techniques for improving website rankings with search engine optimization (SEO). *Adv. Sci. Lett.* **21**(10), 3219–3224 (2015). <https://doi.org/10.1166/asl.2015.6503>
10. Choudhari, K., Bhalla, V.K.: Video search engine optimization using keyword and feature analysis. *Procedia Comput. Sci.* **58**, 691–697 (2015). <https://doi.org/10.1016/j.procs.2015.08.089>
11. Deitel, H.M.: *Internet & World Wide Web*, 4th edn. Pearson (2007)
12. Di Gangi, P.M., Wasko, M.: Social media engagement theory: exploring the influence of user engagement on social media usage. *J. Organ. End User Comput.* **28**(2), 53–73 (2016). <https://doi.org/10.4018/JOEUC.2016040104>
13. Enge, E., Spencer, S., Stricchiola, J.: *The Art of SEO Mastering Search Optimization*, 3rd edn. O'Reilly Media, Inc. (2015)
14. Ghulam, A., Depar, M.H., Ali, S., Rahu, S.: On-Page search engine optimization (SEO) techniques model: a use case scenario of a business. *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)* **5**(11), 3076–3083 (2017)
15. Google: Block Search indexing with “noindex.” (2021a). <https://developers.google.com/search/docs/advanced/crawling/block-indexing>
16. Google: Introduction to robots.txt (2021b). <https://developers.google.com/search/docs/advanced/robots/intro>
17. Google: Qualify your outbound links to Google (2021c). <https://developers.google.com/search/docs/advanced/guidelines/qualify-outbound-links>
18. Google: Robots meta tag, data-nosnippet, and X-Robots-Tag specifications (2021d). https://developers.google.com/search/docs/advanced/robots/robots_meta_tag
19. Gupta, S., Rakesh, N., Thakral, A., Chaudhary, D.K.: Search engine optimization: success factors. In: 2016 4th International Conference on Parallel, Distributed and Grid Computing, PDGC 2016, pp. 17–21 (2016). <https://doi.org/10.1109/PDGC.2016.7913146>

20. Janani., V.D., Haribaabu., V., Sajeev, R.A.: Terminating the spamming links and privacy guaranteed search logs. In: 2013 International Conference on Information Communication and Embedded Systems, ICICES 2013, pp. 394–397 (2013). <https://doi.org/10.1109/ICICES.2013.6508382>
21. Khan, M.N.A., Mahmood, A.: A distinctive approach to obtain higher page rank through search engine optimization. *Sadhana—Acad. Proc. Eng. Sci.* **43**(3) (2018). <https://doi.org/10.1007/s12046-018-0812-3>
22. Khraim, H.S.: The impact of search engine optimization on online advertisement: the case of companies using E-marketing in Jordan. *Am. J. Bus. Manage.* **4**(2), 76–84 (2015). <https://doi.org/10.11634/216796061504676>
23. Killoran, J.B.: How to use search engine optimization techniques to increase website visibility. *IEEE Trans. Prof. Commun.* **56**(1), 50–66 (2013). <https://doi.org/10.1109/TPC.2012.2237255>
24. Lali, M.I.U., Ul Mustafa, R., Saleem, K., Nawaz, M.S., Zia, T., Shahzad, B.: Finding healthcare issues with search engine queries and social network data. *Int. J. Semant. Web Inf. Syst.* **13**(1), 48–62 (2017). <https://doi.org/10.4018/IJSWIS.2017010104>
25. Malaga, R.A.: Search engine optimization—black and white hat approaches. In: *Advances in Computers*, vol. 78, no. 10, pp. 1–39. [https://doi.org/10.1016/s0065-2458\(10\)78001-3](https://doi.org/10.1016/s0065-2458(10)78001-3)
26. Manhas, J.: A study of factors affecting websites page loading speed for efficient web performance. *Int. J. Comput. Sci. Eng.* **3**, 32–35 (2013)
27. Moz: Anchor Text (2021a). <https://moz.com/learn/seo/anchor-text>
28. Moz: Canonicalization (2021b). <https://moz.com/learn/seo/canonicalization>
29. Moz: Domain Authority (2021c). <https://moz.com/learn/seo/domain-authority>
30. Moz: Page Authority (2021d). <http://moz.com/learn/seo/page-authority>
31. Mozilla: HTTP response status codes (2021a). <https://developer.mozilla.org/en-US/docs/Web/HTTP/Status>
32. Mozilla: lang—Language tag syntax (2021b). https://developer.mozilla.org/en-US/docs/Web/HTML/Global_attributes/lang
33. Noruzi, A.: A study of HTML title tag creation behavior of academic web sites. *J. Acad. Librarianship* **33**(4), 501–506 (2007). <https://doi.org/10.1016/j.acalib.2007.03.008>
34. Pažeraitė, A., Repovienė, R.: Content marketing decisions for effective internal communication. *Manage. Organ. Syst. Res.* **79**(1), 117–130 (2018). <https://doi.org/10.1515/mosr-2018-0008>
35. Ramlall, S., Sanders, D., Tewkesbury, G., Ndzi, D.: Can keyword length indicate Web Users’ readiness to purchase. *April*, 1–16
36. Rowley, J.: Understanding digital content marketing. *J. Mark. Manag.* **24**(5–6), 517–540 (2008). <https://doi.org/10.1362/026725708X325977>
37. ScreamingFrog: SEO Spider Tabs (2021). <https://www.screamingfrog.co.uk/seo-spider/user-guide/tabs/>
38. SEMrush: Subdomain vs. Subdirectory (2020). <https://www.semrush.com/blog/subdomain-vs-subdirectory/>
39. SEMrush: Site Audit Crawled Pages Report (2021a). <https://www.semrush.com/kb/543-site-audit-crawled-pages>
40. SEMrush: Site Audit Statistics (2021b). <https://pt.semrush.com/kb/544-site-audit-statistics>
41. Tsuei, H.J., Tsai, W.H., Pan, F.Te, Tzeng, G.H.: Improving search engine optimization (SEO) by using hybrid modified MCDM models. *Artif. Intell. Rev.* **53**(1) (2020). <https://doi.org/10.1007/s10462-018-9644-0>
42. Visser, E.B., Weideman, M.: An empirical study on website usability elements and how they affect search engine optimisation. *SA J. Inf. Manage.* **13**(1), 1–9 (2011). <https://doi.org/10.4102/sajim.v13i1.428>
43. Williams, A.: *SEO 2016 & Beyond*, 5th edn. CreateSpace Independent Publishing Platform (2015)
44. Yalçın, N., Köse, U.: What is search engine optimization: SEO? *Procedia—Soc. Behav. Sci.* **9**(July 2009), 487–493. <https://doi.org/10.1016/j.sbspro.2010.12.185>
45. Zhang, S., Cabage, N.: Search engine optimization: comparison of link building and social sharing. *J. Comput. Inf. Syst.* **57**(2), 148–159 (2017). <https://doi.org/10.1080/08874417.2016.1183447>

46. Ziakis, C., Vlachopoulou, M., Kyrkoudis, T., Karagkiozidou, M.: Important factors for improving Google search rank. *Future Internet* **11**(2) (2019). <https://doi.org/10.3390/fi11020032>
47. Zilincan, J.: Search engine optimization. *Future Internet* **12**(1), 506–510 (2015). <https://doi.org/10.3390/fi12010006>

Data Mining and Big Data—Marketing Data Science

The Influence of Emotions and Score Attributes on Online Ratings in P2P Accommodation: A Data Mining Study



E. Bigné , M. L. Fuentes-Medina , and Sandra Morini-Marrero 

Abstract User-generated content posted on social media reveals users' reviews based on their own assessments. Bookings of peer-to-peer accommodations have increased in the last decade, and research is needed to better understand tourists' experiences. Based on 1 million online reviews posted in Airbnb, this study combined a text mining analysis and statistical analysis of online reviews to evaluate whether the emotions expressed in online reviews and the rentals' attributes explained the online ratings. Our findings suggest that emotions and attributes influence online ratings. However, the type of emotion as well as its intensity affected online ratings differently. Furthermore, online reviews in English had higher average scores than reviews in Spanish. Our findings contribute to a better understanding of the antecedents of online ratings generally.

Keyword Airbnb · Ratings · Sentiment analysis · Content analysis

1 Introduction

In the tourism and hospitality industry, the development of social media has boosted user-generated content (UGC) on a wide range of platforms. UGC contains customers' evaluations of products or services, making it an appropriate and reliable way to understand tourists' feelings and perceptions about their experiences. Peer-to-peer (P2P) rentals provide different services than hotel chains [1], such as living like a local and the opportunity to interact with the host and local community [2]. Online

E. Bigné (✉)

Department of Marketing, Universitat de Valencia, Valencia, Spain
e-mail: enrique.bigne@uv.es

M. L. Fuentes-Medina · S. Morini-Marrero

Department of Economics, Accounting, and Finance, IUDE, Universidad de La Laguna,
San Cristóbal de La Laguna, Spain
e-mail: lfuentes@ull.es

S. Morini-Marrero

e-mail: smorini@ull.es

review content allows researchers to identify P2P guests' experience dimensions and analyze what antecedents trigger their comments and the main outcomes.

Online review ratings are commonly used as a proxy for customer satisfaction [3]. The antecedents of guest satisfaction include emotions and the assessment of the experience's attributes. As [4] noted, when tourists recall their experiences, emotions are the most memorable dimension of their visits. Online guests' textual reviews are now considered an effective means of identifying such emotions, as they have been posted on social media [5]. Further, previous studies have analyzed the relationship between emotions and online ratings [3, 6, 7], while others have studied the relationship between P2P accommodations' attributes and online ratings as a proxy for customer satisfaction [3, 6, 8–11].

A hotel's surrounding environment is known to affect online ratings [12–14]. Thus, it seems reasonable that the surrounding environment of P2P accommodations would also affect guests' experiences. However, to the best of our knowledge, no studies have addressed the impact of location on P2P accommodations' guest ratings. Furthermore, understanding of consumer experiences across different languages and cultures is limited. Some studies have examined the impact of language on online ratings because customers' expectations in the hospitality industry differ between nationalities [15, 16]. However, the effect of language on reviews of P2P accommodations is inconclusive. Thus, while some studies [17–19] found differences in ratings of Airbnb accommodations' attributes according to nationality, another [20] suggested that Airbnb guests' experiences were similar across cultures. Therefore, it is imperative to study the possible cross-cultural tourist experience by examining the online reviews of P2P guests who speak different languages.

Despite efforts to understand the concept of satisfaction, no studies have measured customer satisfaction in an Airbnb setting by integrating emotions with P2P attribute ratings. Thus, our research goal was threefold: to analyze how the emotions in online reviews shape the overall online ratings of P2P accommodations, to determine how the attributes indicated in online reviews affect the overall rating, and to assess to what extent language and destination affect the rating in the online review.

2 Theoretical Background and Hypothesis

Previous literature has explored P2P experience variables that are sorted into three types of variables: antecedent variables, the P2P experience, and outcome variables. This study identified two antecedent variables (accommodations' attributes and emotions or sentiments expressed by guests) from extant literature that play a key role in forming the P2P experience, which leads to the outcome variable (customer satisfaction).

The conceptual model builds upon the model proposed by [21], adopting the two levels of attribute–consequence from Gutman's means-end chain theory [22] and the cognitive, affective, and behavioral (CAB) model [23].

2.1 *Emotions and Ratings*

Previous studies of hotels showed a relationship between sentiment and online ratings [24]. In these studies, a change in customer rating was influenced by reviewing sentiment polarity. In P2P accommodations, previous studies indicated that the ratings given by Airbnb users were consistent with the emotional expressions present in their comments [3, 25, 26]. Our study drew on previous findings and paid special attention to the overall emotion and the intensity of the emotion in online reviews. Expanding upon previous findings, we proposed the following hypotheses:

H1 The overall emotion and emotional intensity of online reviews affect the overall rating of P2P accommodations.

H2 Reviews with high valence lead to higher ratings for P2P accommodations.

2.2 *P2P Attributes and Online Ratings*

Previous studies have shown a relationship between rentals' attributes and the overall online rating of P2P accommodations [7, 9, 11, 27]. The attributes that influenced customer satisfaction most was property location, followed by the host–guest interaction. The remaining attributes, such as home facilities, cleanliness, or decoration, exerted less influence. However, [28] found that tourists did not use compensatory models to rate a service. Therefore, the literature yielded inconclusive results. Based on these findings, we proposed the following hypothesis:

H3 The attributes mentioned in the online review affect the overall rating differently.

2.3 *Location of the P2P Property and Ratings*

The location of a rental property is crucial to its performance [12]. The location affects, among other things, hotel guests' ratings [13], and hotel rates [14]. However, to the best of our knowledge, no studies have analyzed the impact of location on guests' ratings or overall experiences of P2P accommodations. Thus, the attractions surrounding the property might influence its overall rating. Therefore, we hypothesize that:

H4 The attractions surrounding the P2P property affect its online rating.

2.4 Language and Online Ratings

Cultural background might affect how customers assess services. Language can be considered a proxy for cultural background. Indeed, cross-cultural analyses have found differences in the scores given to Airbnb properties' attributes [17–19]. However, another study [20] found a similar number of pleasant Airbnb experiences in three countries with different cultural backgrounds and languages. In light of these inconclusive findings, we proposed the following hypothesis:

H5

Online ratings of P2P accommodations differ by language.

3 Method and Data

The data were retrieved from the Inside Airbnb database. The database contained 1,066,929 reviews of 7 Spanish tourist destinations: Barcelona (BCN), Bilbao (BIL), Donostia (DON), Madrid (MAD), Malaga (MAL), Seville (SEV), and Valencia (VAL).

After retrieving the data, we first performed a sentiment analysis, locating specific emotional words from a defined lexicon in each review (NRC Emotion Lexicon), developed by Mohammad (see details here <http://saifmohammad.com/>) to identify Plutchik's [29] basic emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The sentiment analysis was analyzed through (1) the presence or absence of each emotion and (2) the intensity of each emotion in the review. That is, it considered how many words expressed the same emotion within each review. Further, the sentiment analysis measured the valence of each online review, that is, the net sum of all positive and negative words in the review. If the valence was higher than zero, the polarity of the review was positive.

The attributes mentioned in the reviews were identified using content analysis. This paper analyzed four attribute sets: host, accommodation, location, and environment. To identify if the host was mentioned in the review, we searched for first names. Regarding the accommodations, we aimed to determine whether the reviews mentioned the facilities, comfort, decoration, etc. For this analysis, the location included features of the property's surrounding area, such as its proximity to transportation, leisure areas, and similar places, while the environment referred to the city and its tourist resources.

We tested the hypotheses through multiple statistical analyses, depending on the nature of the variables. The Wilcoxon–Mann–Whitney test allowed us to analyze whether there were differences in a property's rating depending on the presence of each basic emotion in the review, a positive polarity, and whether the review mentioned the four analyzed attributes. This test was also used to assess whether there were differences in ratings based on the review's language.

The Kruskal–Wallis test was used to analyze differences based on the town where the property is located with the support of the Dunn test for post hoc analysis. To measure whether there was a statistically significant relationship between the score, the intensity of the emotions, and the valence of the review, we analyzed the individual significance of the regression coefficient of the model consisting only of the independent variable. Finally, the independent variables were incorporated into a linear regression model to calculate the contribution of each one to the property's online rating. Two models were estimated, one considering the presence of emotions and their positive polarity and the other considering the intensity of emotions and the valence of the reviews.

4 Results and Discussion

Table 1 shows a descriptive summary of the sample, considering all variables involved in the analysis. Trust and joy were the most expressed emotions in the reviews, and disgust was the least expressed. Almost 89% of reviews had a positive polarity. Finally, the attribute relative to the host was mentioned in almost 90% of the reviews.

The average overall score was 9.38 over 10, with a standard deviation of 0.46. The first quartile included scores up to 9.2, and only 2% of the scores were below 5.

Regarding the presence of basic emotions in the reviews and their impact on accommodations' ratings, the Wilcoxon–Mann–Whitney test detected statistically significant differences in the average online ratings given by guests who expressed joy, surprise, trust, or anticipation (p -value $< 2.2e-16$). Reviews with these emotions had higher average scores than reviews that did not. There were also statistically significant differences when the review expressed disgust, anger, fear, or sadness (p -value $< 2.2e-16$). Thus, the average scores of reviews that did not express these emotions were higher.

Table 1 Descriptive data summary (frequencies)

Language and destination		Frequency of emotions		Frequency of attributes	
English reviews	69.03%	Joy	80.65%	Host	89.77%
BCN	21.88%	Surprise	33.58%	Accommodations	67.11%
BIL	2.02%	Trust	87.12%	Location	56.38%
DON	3.31%	Anticipation	60.87%	Destination	20.46%
MAD	33.24%	Disgust	9.48%		
MAL	10.22%	Anger	12.77%		
SEV	16.94%	Fear	15.02%		
VAL	12.39%	Sadness	31.2%		
		Positive polarity	88.89%		

(Source Own elaboration)

A simple linear regression was conducted to determine the influence of an emotion's intensity on online ratings. The results indicated that all emotions showed the expected sign and were significantly related to the overall score (p -values $< 2.2e-16$). In addition, the higher the intensity of positive emotions, the higher the score, and the higher the intensity of negative emotions, the lower the score. Therefore, the results seemed to support hypothesis 1 and were consistent with previous results [3, 25]. Regarding positive polarity and valence, positive polarity generated a statistically significant higher overall score (p -value $< 2.2e-16$), and the valence of the review showed a positive linear relationship with the overall score (p -value $< 2.2e-16$). Thus, the results supported hypothesis 2 and were consistent with the results of previous studies [7].

The univariate analysis of attributes showed that, when the review mentioned the host, the accommodations' characteristics, its location, or the destination's resources, the score was significantly higher (p -value $< 2.2e-16$). Therefore, hypothesis 3 could not be rejected, and the results were consistent with those of previous works [7, 9, 27].

Regarding the influence of the destination, the Kruskal–Wallis test indicated differences between the scores given to accommodations in different destinations. Furthermore, the Dunn test showed such differences between all pairs of cities of the study (p -values $< 2.2e-16$). Therefore, the results seemed to support hypothesis 4 and were consistent with previous studies [13]. There were also significant differences in scores depending on the language of the review: reviews written in English had higher average scores than reviews written in Spanish (p -value $< 2.2e-16$), confirming hypothesis 5.

The joint influence of all variables on the overall online ratings was analyzed with two linear regression models (Table 2). One model considered the presence of emotions in the review and whether the review had a positive polarity (presence model), along with the rest of the variables. The other model considered the intensity of the emotions and the valence of the review (intensity model), along with the rest of the variables. The results of the presence model (Table 2) showed a low coefficient of determination, indicating that, although the analyzed variables influenced the score jointly and individually, they did not capture an important element of score variation. Regarding the attributes, the host, apartment, and destination were always significant. Further, reviews written in Spanish scored lower than reviews written in English, which were used as a reference. Finally, for destinations, Barcelona was the reference, and it showed lower overall scores than other cities. Concerning the intensity model, Table 2 shows that the global adjustment was slightly higher than in the previous model. The intensity of all emotions except sadness was significant.

5 Conclusions

This study analyzed how emotions and the attributes of online reviews shaped the overall online ratings of P2P accommodations. It also assessed the influence of

Table 2 Presence model and intensity model

	Presence model	Intensity model
Intercept	9.069 *	9.196 *
Joy	0.067 *	0.028 *
Surprise	0.033 *	0.008 *
Trust	0.010 *	- 0.005 *
Anticipation	- 0.001	- 0.016 *
Disgust	- 0.076 *	- 0.054 *
Anger	- 0.106 *	- 0.074 *
Fear	0.002	0.006 *
Sadness	0.016 *	0.000
Pos. Pol. Valence	0.122 *	0.022 *
Host	0.073 *	0.063 *
Apartment	0.020 *	0.012 *
Location	0.000	- 0.006 *
Destination	0.022 *	0.002
BIL	0.246 *	0.246 *
DON	0.250 *	0.244 *
MAD	0.082 *	0.083 *
MAL	0.113 *	0.110 *
SEV	0.163 *	0.160 *
VAL	0.109 *	0.107 *
Spanish	- 0.089 *	- 0.091 *
R ²	6.63%	7.75%

(Source Own elaboration)

* Indicates *p*-value < 0.001, and ** indicates *p*-value < 0.05

language and destination on online reviews. Our results showed that emotions significantly influenced online ratings. However, the type of emotion affected online ratings differently. Joy, anticipation, disgust, and anger were significant for all scores. The following attributes mentioned in a post, the host, location of the rental, and resources of the destination, lead to higher overall rating. The attractions surrounding the property affect its online rating. Furthermore, online reviews differed by language. Thus, reviews in English had higher average scores than reviews in Spanish. These findings provide implications for managers in three directions. First, hosts should stress the emotional dimension of the tourist experience. Second, hosts should deliver detailed information about the location and the resources. Lastly, the language is affecting the online ratings.

This study had two main limitations that create avenues for future research. The memorability of the experience and their relationships were not analyzed over time.

References

1. Yi, J. Yuan, G., Yoo, C.: The effect of the perceived risk on the adoption of the sharing economy in the tourism industry: the case of Airbnb. *Inf. Process. Manag.* **5**(1), 1–11 (2020)
2. Tussyadiah, P., Pesonen, J.: Impacts of peer-to-peer accommodation use on travel patterns. *J. Travel Res.* **55**(8), 1022–1040 (2016)
3. Tussyadiah, P., Zach, F.: Identifying salient attributes of peer-to-peer accommodation experience. *J. Travel Tour. Mark.* **34**(5), 636–652 (2017)
4. Kim, J.H., Ritchie, J.R.B.: Cross-cultural validation of a memorable tourism experience scale (MTES). *J. Travel Res.* **53**(3), 323–335 (2014)
5. Rosario, A.B., de Valck, K., Sotgiu, F.: Conceptualizing the electronic word-of-mouth process: what we know and need to know about eWOM creation, exposure, and evaluation. *J. Acad. Mark. Sci.* **48**(3), 422–448 (2020)
6. Subroto, A., Christianis, M.: Rating prediction of peer-to-peer accommodation through attributes and topics from customer review. *J. Big Data* **8**(9) (2021)
7. Zhu, L., Cheng, M., Wong, I.A.: Determinants of peer-to-peer rental rating scores: the case of Airbnb. *Int. J. Contemp. Hosp. Manag.* **31**(9), 3702–3721 (2019)
8. Cheng, M., Jin, X.: What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* **76**, 58–70 (2019)
9. Ju, Y., Back, K.J., Choi, Y., Lee, J.S.: Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *Int. J. Hosp. Manag.* **77**, 342–352 (2019)
10. Thomsen, C.R., Jeong, M.: An analysis of Airbnb online reviews: user experience in 16 U.S. cities. *J. Hosp. Tourism Technol.* **12**(1), 97–111 (2021)
11. Xu, F., La, L., Zhen, F., Lobsang, T., Huang, C.: A data-driven approach to guest experiences and satisfaction in sharing. *J. Travel Tour. Mark.* **36**(4), 484–496 (2019)
12. Yang, Y., Wong, K.K.F., Wang, T.: How do hotels choose their location? Evidence from hotels in Beijing. *Int. J. Hosp. Manag.* **31**(3), 675–685 (2012)
13. Banerjee, S., Chua, A.Y.: In search of patterns among travellers' hotel ratings in TripAdvisor. *Tour. Manage.* **53**, 125–131 (2016)
14. Soler, I.P., Gemar, G., Guzman, V.F.: A comparison of destinations' impacts on hotel rates. *Int. J. Hosp. Manag.* **77**, 226–237 (2019)
15. Armstrong, R.W., Mok, C., Go, F.M., Chan, A.: The importance of cross-cultural expectations in the measurement of service quality perceptions in the hotel industry. *Int. J. Hosp. Manag.* **16**(2), 181–190 (1997)
16. Ariffin, A.A., Maghzi, A.: A preliminary study on customer expectations of hotel hospitality: Influences of personal and hotel factors. *Int. J. Hosp. Manag.* **31**(1), 121–198 (2012)
17. Cheng, M., Zhang, G.: When Western hosts meet Eastern guests: Airbnb hosts' experience with Chinese outbound tourists. *Ann. Tour. Res.* **75**, 288–303 (2019)
18. Xi, Y., Ma C., Yang, Q., Jiang, Y.: A Cross-cultural analysis of tourists' perceptions of Airbnb attributes. *Int. J. Hosp. Tourism Adm.* 1–34 (2021). <https://www.tandfonline.com/doi/full/10.1080/15256480.2020.1862014>
19. Zhu, Y., Ma, L., Jiang, R.: A cross-cultural study of English and Chinese online platform reviews: a genre-based view. *Discourse Commun.* **13**(3), 342–365 (2019)
20. Brochado, A., Troilo, M., Shah, A.: Airbnb customer experience: evidence of convergence across three countries. *Ann. Tour. Res.* **63**, 210–212 (2017)
21. Zhang, H., Wu, Y., Buhalis, D.: A model of perceived image, memorable tourism experiences and revisit intention. *J. Destin. Mark. Manag.* **8**(June), 326–336 (2018)
22. Gutman, J.: A means-end chain model based on consumer categorization processes. *J. Mark.* **46**, 60–72 (1982)
23. Lavidge, R.C., Steiner, G.A.: A model for predictive measurements of advertising effectiveness. *J. Mark.* **25**, 59–62 (1961)
24. Geetha, M., Singha, P., Sinha, S.: Relationship between customer sentiment and online customer ratings for hotels—an empirical analysis. *Tour. Manage.* **61**, 43–54 (2017)

25. He, W., Tian, X., Tao, R., Zhang, W., Yan, G., Akula, V.: Application of social media analytics: a case of analyzing online hotel reviews. *Online Inf. Rev.* **41**(7), 921–935 (2017)
26. Meijerink, J., Schoenmakers, E.: Are online reviews in the sharing economy skewed toward positive ratings? Linking customer perceptions of service quality to leaving a review of an AirBnB stay. *J. Tourism Futures* **7**(1), 5–19 (2021)
27. Joseph, G., Varghese, V.: Analyzing Airbnb customer experience feedback using text mining. In: Sigala M., Rahimi R., Thelwall M. (eds) *Big Data and Innovation in Tourism. Travel, and Hospitality*, pp. 147–162. Springer, Singapore (2019).
28. Bigné, E., William, E., Soria-Olivas, E.: Similarity and consistency in hotel online ratings across platforms. *J. Travel Res.* **59**(4), 742–758 (2020)
29. Plutchik, R.: A general psychoevolutionary theory of emotion. In: Plutchik, R., Kellerman, H. (eds.) *Emotion: Theory, Research, and Experience*, vol. 1, pp. 3–31. Academic Press, New York (1980)

Sensing the Impact of COVID-19 Restrictions from Online Reviews: The Cases of London and Paris Unveiled Through Text Mining



Bruno Silva, Sérgio Moro, and Catarina Marques

Abstract This study aims to understand how the COVID-19 pandemic affected the hotel sector and to identify the current traveler demands. The traveler's reviews were analyzed based on sentiment analysis and a guest satisfaction model was also proposed, demonstrating a data mining approach within tourism and hospitality research. Given its popularity, TripAdvisor was the chosen platform for collection of hotel reviews in London and Paris. Text data were extracted from reviews made in two time periods, before and during the COVID-19 pandemic. The sentiment and specific aspects highlighted by travelers were compared between each period.

Keywords Text mining · Sentiment analysis · Tourism · Hotel traveler's online reviews · COVID-19 pandemic

1 Introduction

Over the past years, we have witnessed a notorious increase in the number of text reviews made in several online accommodation platforms following the exponential growth in tourism [3]. However, during the year of 2020 as the COVID-19 pandemic continued to spread, the online reviews amount has completely changed, namely with the drastic decrease of travel research [7].

COVID-19 pandemic brought a radical change in people's lives with many economic sectors severely impacted. Tourism was one of the most affected business sectors contrasting with the steady growth seen over the years [3]. For this reason, it is now extremely important to help tourism, and more specifically hotels, to understand how COVID-19 affected the sentiment of travelers during their stays. The present

B. Silva (✉) · S. Moro · C. Marques
ISCTE—University Institute of Lisbon,, Lisbon, Portugal
e-mail: brmsa@iscte-iul.pt

S. Moro
e-mail: sergio.moro@iscte-iul.pt

C. Marques
e-mail: catarina.marques@iscte-iul.pt

research aims to identify a gap on subareas within tourism on which data science has not yet been applied. More particularly, this investigation will focus on data mining on online reviews from the TripAdvisor platform, written during COVID-19 pandemic, to understand what travelers are seeking and what demands they now have. The TripAdvisor reviews were collected from two capital cities, London and Paris, for being very similar in terms of tourism characteristics, focusing on findings answers essentially for one main question: “How is the effect of the pandemic being perceived by visitors to major European capitals severely affected by COVID-19?”

2 Literature Review

2.1 Sentiment Analysis and the Impact of COVID-19

Sentiment analysis is a natural language processing tool that is useful for monitoring Web 2.0 applications, as it can reveal public opinion about numerous issues without requiring satisfaction surveys [8]. The interest in sentiment analysis has increased significantly over the last few years due to the large amount of text stored in Web 2.0 applications and the importance of online customer opinions [8]. As a result, more than 1 million research papers contain the term “sentiment analysis” and various start-ups have been created to analyze sentiments in social media companies [8].

The spread of the COVID-19 global pandemic has generated an exponentially mounting and volume of data that can be harnessed to improve our understanding of big data management research. It is also applied in the need among scholars, practitioners and policymakers for a better and deeper understanding of a range of analytical tools that could be utilized to better anticipate and respond to such unforeseen ‘black swan’ events and risks. [6].

2.2 Conceptual Framing and Research Hypothesis

Based on other studies that identified a list of dimensions drawn from quantitative features known to influence customer satisfaction under the context of TripAdvisor, the list below (Table 1) was considered, as well as a proposal for a guest satisfaction model to illustrate a data mining approach within tourism and hospitality research.

Restuputri et al. [5] published a research comparing the staff service quality, operational, technical logistics service providers, with customer satisfaction and loyalty during the COVID-19 pandemic. In this study, they refer that an employee must be reliable, punctual and careful at work. An employee must also have effective communication skills, be courteous and ready to serve. The quality of operations service from source to customers must be well-coordinated, on time and with appropriate transportation capacity so no damage occurs to the customer’s property. In the end, he

Table 1 List of considered dimensions for this study

Dimension	Reference
Services	Chen et al. [3]
Amenities	Moro et al. [4]
Health measures	Chen et al. [3]
Hotel facilities	Moro et al. [4]
Location	Chen et al. [3]
Value/price	Chen et al. [3]
Cleanliness	Moro et al. [4]
Type of travel	Moro et al. [4]
Hotel prestige	Moro et al. [4]
Seasonality	Moro et al. [4]

concluded that a good relationship between staff, operational and technical services is crucial for a good perception of customer satisfaction and loyalty to the services provided by hotels.

Based on the above, the main hypotheses tested in this research was:

H1 Hotel guests appreciate safety concerns by hotels and express about them in online reviews.

Big crisis affects the macro-environment that brings big changes in customer behaviors and hotel performances [3]. The outbreak of COVID-19 has brought changes in guest experiences at hotels as customers usually describe their experiences and feedbacks after their stays in the form of user-generated content, such as online reviews which are then used by practitioners to understand the nature and structure of guest experiences [3]. Since hotels are customer-centric that should keep up with customer preferences and requirements, it is essential to understand the impact of COVID-19 and track the changes of guest experiences brought by the same [3].

3 Methodology

The aim of this investigation is to propose a guest satisfaction model by seeking any available data and to illustrate a data mining approach within tourism and hospitality research. One of the methods utilized was based on the identification and analysis of a list of specific dimensions drawn from quantitative features known to influence customer satisfaction under the context of TripAdvisor [4].

The method used for this investigation is based on a problem resolution paradigm that contains several activities [1]:

Identification and motivation of problem was the first step which defines that the artefact must be viable, that is, it must provide a solution to the research problem. In

the presented case, it will be produced in the form of a method to study the impact of COVID-19 restrictions on the TripAdvisor reviews from 2 foreign locations very similar in terms of tourism characteristics. Chosen the platform for the data extraction, it was necessary to define the solution objectives, which must be achievable, in addition of the need to understand if this analysis could identify a research gap in the utilization of data mining in tourism.

The third step, Development, involved the current creation of the artefact defined in the first step, that is, the development of a method that allows the analysis and comparison of both hotel reviews. For the sake of the extraction, a pre-existent coding script was chosen to perform the scraping that contained several features to quickly setup and extract data from a website. In this phase, it was necessary to understand if the objective and its influencing factors could be translated into data features and instances [4]. As soon as this was achieved, the data needed also to be prepared (data cleaning) for model training.

Next to development was the Demonstration, which was used to verify the effectiveness of the solution to the problem identified. To achieve this goal, graph plots and word clouds or any other most appropriate visualization methods were used in addition to and a pre-build script to plot the most common words from the hotel reviews.

The qualitative measurement of how well the created artefact supports the solution to the problem corresponded to the next step, the Evaluation, which involved the use of the model previously trained to make predictions versus observations [4].

And lastly, the communication of the results, whose purpose helped highlight and identify the importance of the problem through the dissemination of the developed artefact, that is, its usefulness and its relevance to other researchers.

4 Results

The discussion of the output results will help clarify of what were the major differences within the customer reviews between the pre and during COVID-19 restrictions on equivalent time periods. In addition, a comparison between capital cities is provided.

Figure 1 represents the number of reviews extracted per year and city. A significant reduction can be immediately seen from 2019 to 2020 in both cities. In case of London, there is, in average, a reduction of 86% on the reviews and Paris with a 95% of decrease. Nevertheless, on 2020, London had a slight increase between June and October mainly due to ease of lockdown restrictions in that period (Institute for Government 2019), followed then by a second decrease as the number of deaths from COVID-19 disease were rapidly rising, as per Fig. 2.

Figure 2 provides a more detailed insight about the evolution of reviews of TripAdvisor and the number of COVID-19 cases and deaths, reported in each city per month. The data from 2019 is also showed to have a fair comparison in the homologous period. In 2020, an increase in the number of reviews, in London, can be seen during

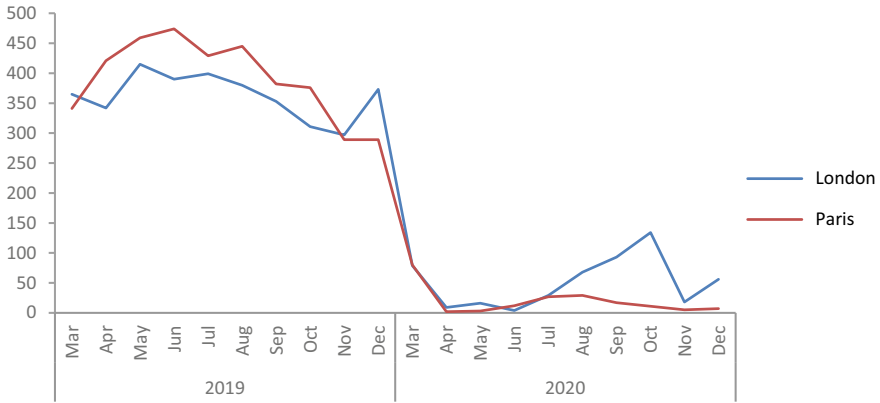


Fig. 1 Distribution of reviews during 2019/2020 periods

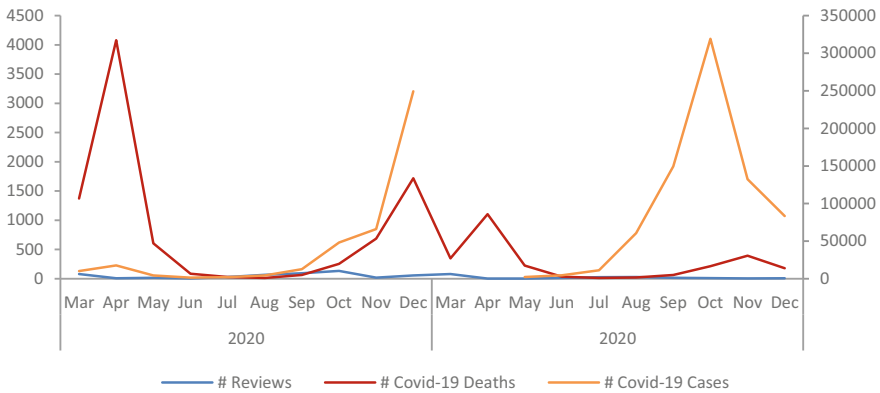


Fig. 2 Distribution of reviews and COVID-19 cases/deaths during the 2020 period

the peak of Summer. An increase is also noticeable in Paris but with much less evidence. This difference is related to the severity of the restrictions implemented on each city which were higher in Paris.

In fact, this chart allows to determine the periods where each government applied and eased the pandemic lockdown restrictions. In the case of London, the ease of restrictions started on June 23, 2020 and on October 31, 2020, England announces second national lockdown (Institute for Government 2019). During this period, the number of reviews done on TripAdvisor rose compared with the first wave of cases.

On a same analogy, in Paris, the government eased the restrictions on June 14, 2020. On August 20, France announced that the pandemic was again on the rise [2]. Due to there was also a slight increase in TripAdvisor reviews but with much less expression, compared to London.

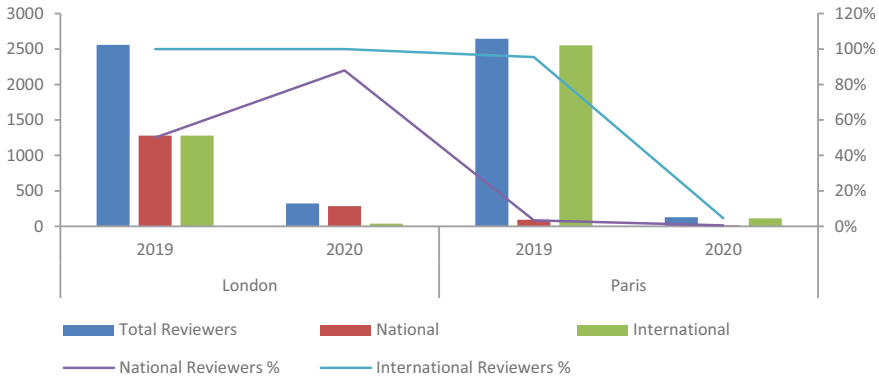


Fig. 3 Comparison between national and international reviewers per city and year

Figure 3 shows the difference in terms of the nationality of the TripAdvisor reviewers. The total number of reviewers categorized by nationality and the percentage based on the total in each year are compared. The scenario of London, in 2020, is an example of expected decrease in the total number of reviewers (87%) and when comparing in terms of nationality a significant increase is noticeable on the national tourists of about 40%.

Paris, on the other hand, demonstrates that the city is more visited by the international tourists than those living within borders, in 2020. This behavior could be explained by the fact of the collected data being filtered in the English language only. Also, this last point can explain the significant decrease in the written reviews on TripAdvisor (95%) as it is very likely that, for Paris, there are even more reviews written in French.

The sentiment analysis part was also included in the mining process of TripAdvisor’s extracted data. Figures 4, 5 and 6 present three different perspectives on the analysis of guest satisfaction based on the dimensions chosen for the study that were already discussed in previous chapters.

To evaluate the guest satisfaction, in each review, the sentiment words were drawn, and their scores were calculated. A score higher than zero was considered a positive review, a score equal to zero a neutral score and finally a score less than zero a negative one. For example, reviews with the words “helpful,” “greeting” and “friendly,” “amazing” got the best scores and reviews with word combinations like “excellent space optimization” and “real pleasant surprise” also were classified with best positive scores. After the identification of the key words for each review, a new classification was done using the list of ten dimensions shared in the study. An average was then calculated per each dimension based on positive and negative perspectives.

As per Figs. 4 and 5, almost every dimension was classified with a positive sentiment, in both years. It is possible to see that in Paris during 2020, seasonality negatively impacted the reviewers when scoring their stays. This may be justified by the pandemic situation lived during this time. Health measures is another dimension

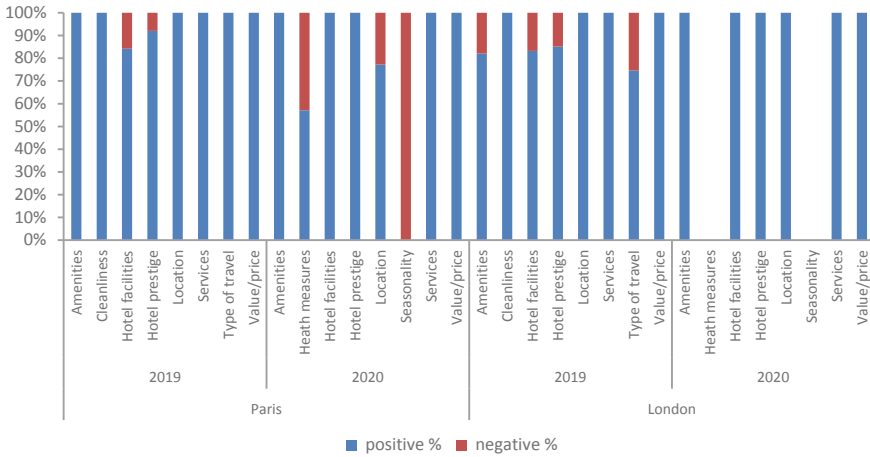


Fig. 4 Sentiment perception distribution by dimension over total number of reviews

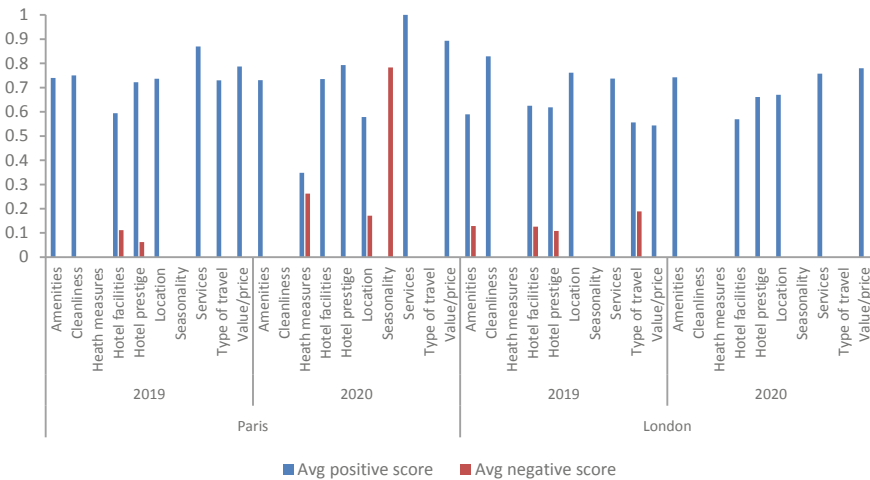


Fig. 5 Sentiment perception average score by dimension, city and year over total number of reviews

affected by this reason getting 42% of negative reviews. In the case of London, curiously, was not possible to retrieve any feedback referring both seasonality and health measures dimensions.

Figure 6 shows another perspective of the dimension analysis. As already referred, seasonality clearly stands out as the most negative perceived dimension in 2020. Health measures were also classified with low review scores. In 2019, the dimensions with poorer review sentiment were related to amenities, hotel facilities and prestige and the type of travel of the guests.

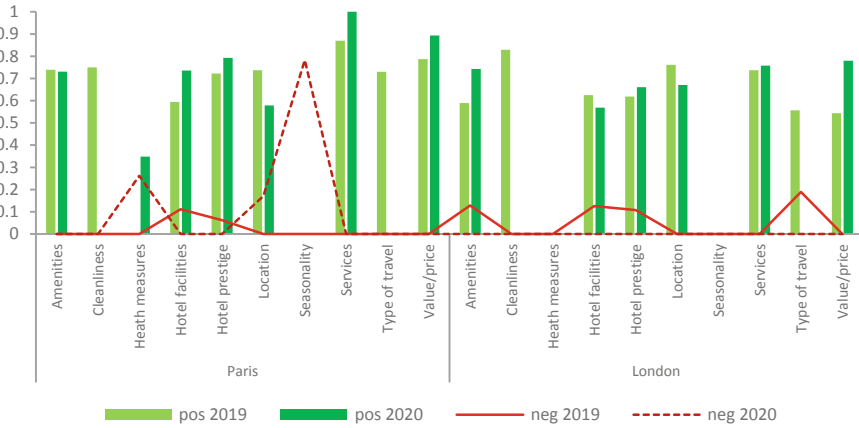


Fig. 6 Sentiment perception by dimension and year over total number of reviews

In the other hand, dimensions like the services and value/price were the ones with better sentiment classification, in both cases of Paris and London.

Table 2 contains a summary of the analysis of each of the dimensions discussed.

5 Conclusion

The main objective of this study was to understand what travelers are seeking nowadays during the current pandemic that we are all living. From the results and analysis done on the data mining process, it was possible to understand some aspects of the behavior, profile and most importantly how guests are now perceiving the effect of COVID-19 pandemic in their hotel stays in two major capital cities being this the major research question of this investigation.

A guest satisfaction method was proposed demonstrating that most of the reviews (55%) done on TripAdvisor on 2020 showed a positive perception, despite the COVID-19 pandemic when comparing to 2019.

Some limitations were also felt during the analysis and data extraction that could not allow for a deeper investigation on some interesting hypothesis. The need of travel insurance has raised with the pandemic or the existence of a full refund option in case of travel cancelling could be some examples. To achieve this, it could be worth to perform a deeper analysis into the TripAdvisor reviews data such as extracting the day guests book their travel. Despite not being a mandatory field, it could lead to the testing of other hypotheses such as if the fear of confinement leads visitors to book their travels only on the same day, they will actually travel.

Table 2 Effects on customer satisfaction

Dimension	Paris	London
Services	In general, services have a positive effect on the guests; 2019—the best rated dimension; 2020—the worst rated	The dimension has also a positive perceived value on customer; 2019—best rated than in 2020 (with a neutral feedback)
Amenities	74% of positive feedback in both years. No negative effect perceived by guests	2020—the second-best rated dimension; 2019—the second worst rated dimension. Nevertheless, the average positive rating is greater than the negative feedback
Health measures	2019—top negative rated dimension; 2020—second negative rated dimension along with Cleanliness	No feedback extracted
Hotel facilities	Third negative dimension in 2019; Mostly rated positive in 2020	Positive rating in general
Location	Increase of 17% of negative rating value in 2020 compared to 2019	100% positive feedback
Value/price	Top positive rated dimension in 2020 and in overall	Second-best positive rated dimension
Cleanliness	Second most negative rated dimension	Top positive rated dimension in 2019
Type of travel	More decisive for travelling in 2019	More decisive for travelling in 2019
Hotel prestige	Mostly rated positively in 2020	Positive rating improvement from 2019 to 2020
Seasonality	Mostly negative perceived by guests in 2020	No feedback retrieved from analyzed guests reviews

References

- Alexandre, I.M.: Design Science Research. 1–19 (2017)
- Aurore, J., Blue, F.: Coronavirus: the key dates of the epidemic in France. (2020). <https://www.francebleu.fr/infos/societe/coronavirus-les-dates-cles-de-l-epidemie-en-france-1603646805>
- Chan, I.C.C., Ma, J., Ye, H., Law, R.: A comparison of hotel guest experience before and during pandemic: evidence from online reviews. In: Information and Communication Technologies in Tourism 2021, pp. 549–556. Springer International Publishing. https://doi.org/10.1007/978-3-030-65785-7_52
- Moro, S., Esmerado, J., Ramos, P., Alturas, B.: Evaluating a guest satisfaction model through data mining. *Int. J. Contemp. Hosp. Manag.* **32**(4), 1523–1538 (2019). <https://doi.org/10.1108/IJCHM-03-2019-0280>
- Restuputri, D.P., Indriani, T.R., Masudin, I.: The effect of logistic service quality on customer satisfaction and loyalty using kansei engineering during the COVID-19 pandemic. <http://www.Editorialmanager.Com/Cogentbusiness> **8**(1), 1906492 (2021). <https://doi.org/10.1080/23311975.2021.1906492>

6. Sheng, J., Amankwah-Amoah, J., Khan, Z., Wang, X.: COVID-19 pandemic in the new era of big data analytics: methodological innovations and future research directions. *Br. J. Manag.* **1467–8551**, 12441 (2020). <https://doi.org/10.1111/1467-8551.12441>
7. Uğur, N.G., Akbıyık, A.: Impacts of COVID-19 on global tourism industry: a cross-regional comparison. *Tourism Manag. Perspect.* **36**, 100744 (2020). <https://doi.org/10.1016/j.tmp.2020.100744>
8. Valdivia, A., Luzón, M.V., Herrera, F.: Sentiment analysis in TripAdvisor. *IEEE Intell. Syst.* **32**(4), 72–77 (2017). <https://doi.org/10.1109/MIS.2017.3121555>

Sources Used to Data-Driven Decision-Making in Tourism Management. Identifying the Main Areas of Research



Juan Vidal, Ramón A. Carrasco , María F. Blasco, and Manuel J. Cobo

Abstract The tourism sector is one of the most affected by the health situation caused by COVID-19. As a result, digital transformation is accelerating in this sector. One of the pillars of this transformation is the management of organizations based on data-driven decision-making. The raw material for such data-driven strategies is obviously the sources of information used. This paper attempts to give a knowledge map of the diverse sources of information used in tourism for this decision-making. To this purpose, we analyse the scientific publications of the last five years in order to identify the main areas of action related to the sources used for data-driven management in tourism. As a result of this bibliometric analysis, we have identified 14 topics that have attracted the interest of the scientific community grouped into three main areas of action.

Keywords Data-driven decision-making · Tourism management · Tourism data sources

1 Introduction

Data-driven decision-making is an area of crucial importance in the digital transformation in which society in general and many organizations in particular are immersed. These decision-making methodologies allow organizations to be truly market oriented, enabling them to focus on customers in order to build customer loyalty in a more cost-effective way than their competitors [1].

Of course, the tourism sector is no exception to this situation. Moreover, this sector has been one of the worst affected by the health situation caused by COVID-19 and is considering a deep transformation in which this type of management will be the predominant one.

J. Vidal · R. A. Carrasco (✉) · M. F. Blasco
Faculty of Commerce and Tourism, Complutense University of Madrid, 28003 Madrid, Spain
e-mail: ramoncar@ucm.es

M. J. Cobo
Department of Computer Science and Engineering, University of Cádiz, 11519 Cádiz, Spain

The raw material for such data-driven strategies is obviously the sources of information used for decision-making. Therefore, it would be interesting to have a knowledge map of the diverse sources of information used in tourism for this decision-making. The aim of this paper is to respond to this knowledge by means of a bibliometric analysis based on the scientific publications of the last five years that allows us to identify the main areas of action related to the sources used for data-driven management in tourism.

To this purpose, Sect. 2 shows the methodology used for the development of this work. Section 3 applies the methodology to identify the areas mentioned above. Finally, conclusions and future work are suggested.

2 Methodology of the Thematic Analysis

2.1 Methodology and Tools Used

The study to be carried out can be classified as a bibliometric analysis. This term was first introduced by Pritchard [2]. The leading bibliometric methods are performance analysis, which focuses on scientific impact and the number of citations reached, and scientific mapping, which refers to the graphical representation of the structure of scientific research in the intellectual, theoretical or social sphere [3]. For this type of study, it is common to use co-word analysis which is a powerful bibliometric technique for identifying, describing and visualizing the interactions between keywords, terms and topics. More precisely, it evaluates the frequency of co-occurrence, i.e. the number of documents in which two keywords appear together [4]. Therefore, the analysis of co-occurrences is a step prior to the construction of the strategy diagrams. Such diagrams are very powerful as a tool for analysis as they allow the semantics to be given to the different issues identified (see Fig. 1).

However, in many cases, themes have logical groupings that can only be identified by experts in the field of knowledge through a more qualitative analysis. Combining these quantitative and qualitative techniques, the methodology shown in Fig. 2 is proposed in this paper.

There are a multitude of bibliometric analyses, but if we restrict ourselves to studies related to the one proposed in this article, we would have the following ones: Kirtil and Askun [6] and Lv et al. [7], although these works are oriented to the use of artificial intelligence and/or big data in tourism and not specifically to the data sources used for data-driven management in tourism.

In this research, we used the SciMAT software (science mapping analysis software tool) [8] to perform the bibliometric analysis, since it integrates everything we required for this research and because it is free and open-source software.

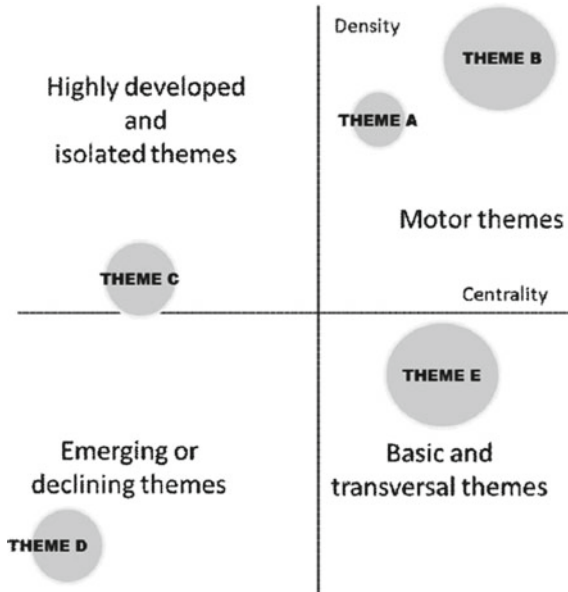


Fig. 1 Example of strategic diagram and thematic network [4]

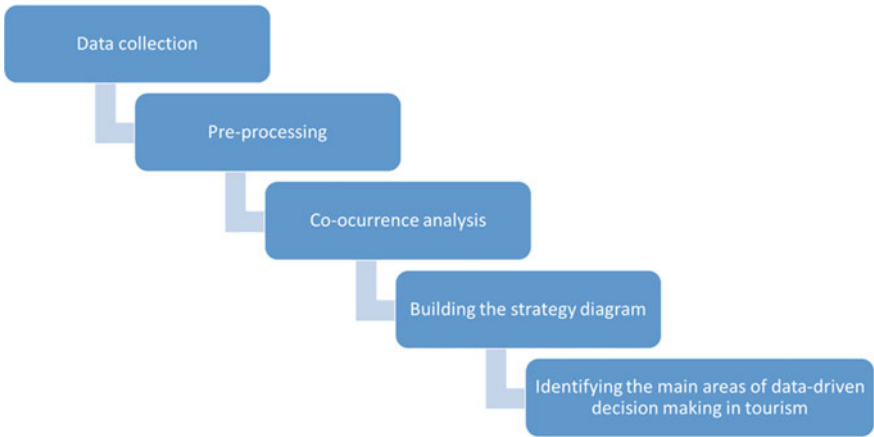


Fig. 2 Methodology of the thematic analysis

2.2 Data Set

Although there are a variety of scientific databases such as Google Scholar and Scopus, the Web of Science (WoS) database can be seen as a standard in the scientific field [1]. This database is therefore the one most commonly used for bibliometric analysis [5]. Therefore, we will extract the data from our study of WoS which includes, among other things, the keywords provided by the authors of the papers.

The time period analysed is from 2016 to 2020.

3 Bibliometric Thematic Analysis

In order to achieve the thematic bibliometric analysis objective of this work, we will follow the steps shown in Fig. 2, which are detailed in the following sections.

3.1 Data Collection

Bibliographic records have been downloaded from the main WoS collection for period 2016–2020, based on the following query:

$TS = (("tourism" OR "tourist*" OR "hospitality" OR "hotel*") NEAR/3 ("data source*" OR "information source*" OR "dataset*" OR "information system*" OR "database*" OR "data warehouse*" OR "data lake*"))$.

Here, the TS field is a search based on a given topic (title, abstract and keywords).

A total of 328 documents were obtained. The data were downloaded on 17/05/2021. The documents were exported from WoS as plain text file including the complete record. This file was imported into SciMAT to build the database from which the bibliometric analysis will be performed. To improve the quality of the data, a debugging process was performed in which words representing the same concepts were grouped together.

3.2 Pre-processing

To improve the quality of the data, a debugging process was performed in which words representing the same concepts were grouped together.

3.3 *Co-occurrence Analysis*

A co-word analysis [9] is applied to the raw data. Next, we performed keyword clustering using the simple centre algorithm [10] that identify networks of keywords that are closely linked to each other and correspond to centres of interest or research problems of significant interest among researchers. This similarity between keywords is evaluated using the equivalence index [10]: $e_{ij} = c^2_{ij}/c_i c_j$, where c_{ij} is the number of documents in which two keywords i and j coexist and c_i and c_j represent the number of documents in which each appears.

3.4 *Building the Strategy Diagram*

Once the co-word analysis has been carried out, the strategy diagram shown in Fig. 3 is constructed. In this strategic diagrams, the size of the sphere of each research topic is proportional to the number of papers published on that topic.

Next, we obtain the performance analysis. In this phase, the contribution of research topics and thematic areas to the entire research field is measured, both quantitatively and qualitatively, to establish which are the subfields with the greatest importance, production and impact. The performance analysis is performed as a complement to the scientific mapping work, and for this purpose the following bibliometric indicators are used: number of published papers, number of citations and H-index [11]. To evaluate these performance indicators, the SciMAT programme uses a function that assigns a set of documents to each detected topic. SciMAT obtains the algebraic union of the set of documents associated with the keywords of the topic. A document has several keywords, since each of them can be associated with a different topic, so that a document could be associated with several topics.

In Table 1, we can see the performance indicators: number of documents, citations obtained by these documents and the H-index. According to these indicators, we could highlight the following as the main topics: determinants, tourism destination, online reviews, management, impact, machine learning and information sources. These seven topics together account for 83% of the total number of citations and 81% of the documents. The H-index of these topics is also higher, although there is not a great difference as in the number of citations or documents.

3.5 *Identifying the Main Areas of Data-Driven Decision-Making in Tourism*

The main themes identified in the previous step are as follows:

- **Online review:** It is a motor theme; it is the third theme with the highest number of documents and the second in terms of citations, equal with determinants theme. It

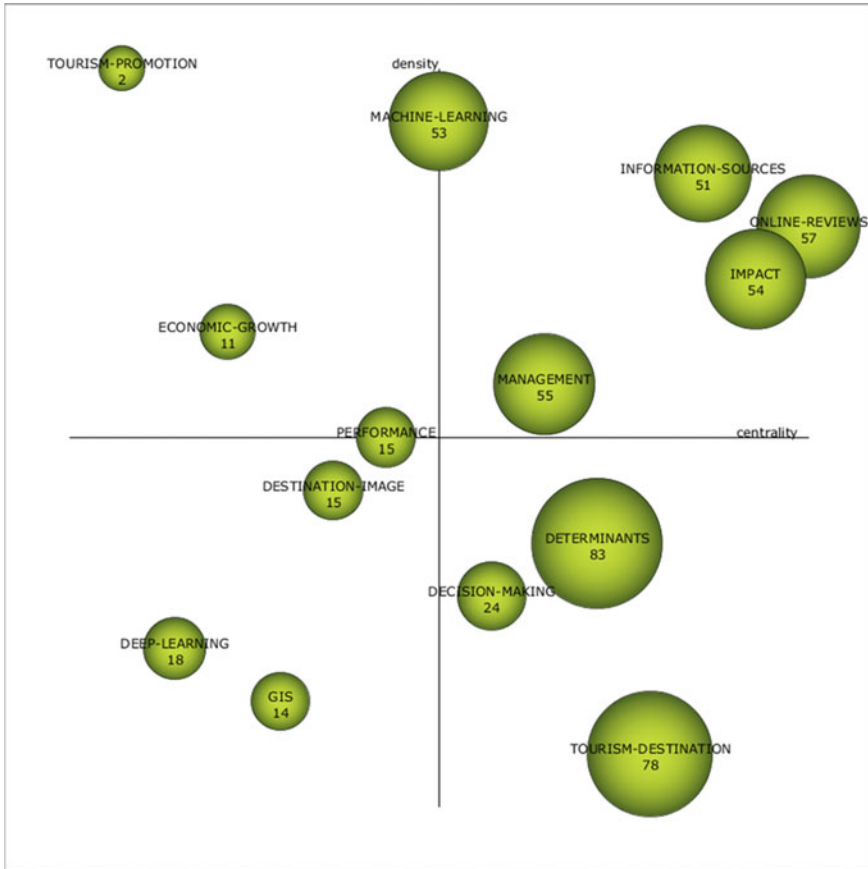


Fig. 3 Strategic diagram for the period

focuses on the information generated by the tourist, either through evaluations or opinions on platforms or marketplaces of tourism services, satisfaction surveys, comments on social networks, exchange of information in virtual communities or e-mails. It can be classified into two subgroups: electronic word of mouth (eWOM) and user-generated content (UGC). Within this theme, there are keywords related to data analysis. It reflects the great relevance that the analysis of tourist-generated information has acquired for tourism, both because technology makes it possible to capture and store it and because of the shift towards a more focused approach to analysing tourist behaviour and experience.

- **Information sources:** It is a motor theme; it is the third theme with the highest number of citations, and its importance is minor in terms of number of documents or H-index. It includes topics related to the different data sources differentiated by their type and origin. It covers data from travel, Internet search engines, images,

Table 1 Themes performance indicators

Theme	Documents	H-index	Citations
Determinants	83	14	705
Tourism destination	78	15	1127
Online reviews	57	14	846
Management	55	8	379
Impact	54	13	627
Machine learning	53	8	481
Information sources	51	13	725
Decision-making	24	7	205
Deep learning	18	7	228
Performance	15	4	48
Destination image	15	5	140
GIS	14	6	169
Economic growth	11	5	168
Tourism promotion	2	1	2

tourist behaviour or hotel reservations. It reflects the wide range of available data sources that has been growing in variety over the last years.

- **Impact:** It is a motor theme; it addresses differentiated topics, including the impact of social networks on tourism, which are becoming increasingly relevant, and on the other hand, a classic topic such as the impact of tourism on the economy, with emphasis on macroeconomics. On the other hand, it deals with issues related to the hospitality industry, distribution channels and online tourist agencies.
- **Management:** It is a motor theme; of the main themes, it is one of the least important in terms of number of documents, number of citations or H-index. It contains several topics, including tourism information systems and organizational issues such as management and service quality.
- **Machine learning:** It is a motor theme; similarly to the previous theme, compared to the rest of the main themes, it is one of the least important in terms of number of papers, number of citations or H-index. It focuses on data analysis techniques applied in tourism research such as machine learning, recommenders, classification algorithms, collaborative filtering or variable selection. It also refers to more general topics, but also related to data analysis such as big data or data mining.
- **Tourism destinations:** It is a basic and transversal theme; it is the one with the highest number of citations, the highest H-index and the second highest number of documents. It focuses on analysing tourism destinations from the perspective of demand, tourist behaviour and experience. In this theme, we find keywords related to current issues such as sustainable tourism and smart tourism that relies on technology to improve services and the tourist experience. The analysis of tourism destinations from different perspectives is one of the central themes in data-driven tourism research.

- **Determinants:** It is a basic and transversal theme; it is the theme with the highest number of papers, the second highest number of citations and also the second highest H-index equal with online reviews. It focuses on the main factors that determine tourism demand. We also find a wide variety of issues related to hotels, the industry, trust, differentiation and prices.

Based on this study, we can detect the following three major areas of research in data-driven analysis in tourism sector:

- **Research focused on analysing tourist behaviour** at a descriptive or predictive level with an increasing use of advanced analytical techniques. Tourist decision-making is increasingly based on more factors and is heavily supported by information generated by other users on platforms or social networks. This information is increasingly analysed to understand tourist behaviour. In this research area, we also find researches on tourism demand forecasting.
- **Research at an organizational level focused on tourism management issues** such as marketing actions or the analysis of tourist destinations from different perspectives such as sustainable tourism or smart tourism (use of technology to improve quality and experience).
- **Research aimed at performance, quality and impact measures in tourism sector:** economic performance of the different actors in the tourism industry such as hotels, travel agencies and others, measures of the quality of tourism service and the measurement of the impact of tourism on macroeconomics

4 Conclusions and Future Work

In this work, a bibliometric analysis of tourism research based on data sources has been carried out. WoS publications from 2016 to 2020 have been used. It consisted of a performance analysis with the most important indicators and a thematic analysis through science maps. The following conclusions can be obtained:

As a result of the bibliometric analysis, we have identified 14 topics that have attracted the interest of the scientific community: determinants, tourism destination, online reviews, management, impact, machine learning, information sources, decision-making, deep learning, performance, destination image, GIS, economic growth and tourism promotion.

The following major areas of research in data-driven analytics have been identified:

- Research focused on the analysis of tourist behaviour at a descriptive or predictive level with an increasing use of advanced analytics techniques.
- Research at the level of organizations focused on tourism management issues such as marketing actions or the analysis of tourist destinations.
- Research oriented to performance and impact measurements in the tourism sector.

Future work should be aimed at analysing the temporal evolution of the topics taking a broader time frame. On the other hand, our study shows the importance of information sources and data analysis techniques in data-driven tourism research. A broader bibliometric analysis could address these issues. Similarly, the use of databases other than WoS may provide a broader view.

References

1. Moreno, C., González, R.A.C., Viedma, E.H.: Data and artificial intelligence strategy: a conceptual enterprise big data cloud architecture to enable market-oriented organisations. *IJIMAI* **5**(6), 7–14 (2019)
2. Pritchard, A.: Statistical bibliography or bibliometrics. *J. Document.* **25**(4), 348–349 (1969)
3. Montero, J., Cobo, M.J., Gutiérrez, M., Segado, F., Herrera, E.: A science mapping analysis of Communication WoS subject category (1980–2013). *Comunicar* **26**(55), 81–91 (2018)
4. del Barrio-García, S., Muñoz-Leiva, F., Golden, L.: A review of comparative advertising research 1975–2018: thematic and citation analyses. *J. Bus. Res.* **121**, 73–84 (2020)
5. Rodríguez-López, M.E., Alcántara-Pilar, J.M., Del Barrio-García, S., Muñoz-Leiva, F.: A review of restaurant research in the last two decades: a bibliometric analysis. *Int. J. Hosp. Manage.* **87**(102387), 1–10 (2020)
6. Kirtil, I.G., Askun, V.: Artificial intelligence in tourism: a review and bibliometrics research. *Adv. Hospital. Tour. Res.-AHTR* **9**(1), 205–233 (2021)
7. Lv, H., Shi, S., Gursoy, D.: A look back and a leap forward: a review and synthesis of big data and artificial intelligence literature in hospitality and tourism. *J. Hospital. Mark. Manage.* (2021)
8. Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., Herrera, F.: SciMAT: a new science mapping analysis software tool. *J. Am. Soc. Inform. Sci. Technol.* **63**(8), 1609–1630 (2012)
9. Callon, M., Courtial, J.P., Turner, W.A., Bauin, S.: From translations to problematic networks: An introduction to co-word analysis. *Soc. Sci. Inf.* **22**(2), 191–235 (1983)
10. Coulter, N., Monarch, I., Konda, S.: Software engineering as seen through its research literature: a study in co-word analysis. *J. Am. Soc. Inf. Sci.* **49**(13), 1206–1223 (1998)
11. Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E., Herrera, F.: H-Index: a review focused in its variants, computation and standardization for different scientific fields. *J. Inform.* **3**(4), 273–289 (2009)

Assessing Corporate Reputation from Online Employee Reviews



R. E. Loke and IJ. A. A. Steentjes

Abstract The overarching aim of this paper is to define, develop and present a processing pipeline that has practical application for companies, meaning, being extendable, representative from marketing perspective, and reusable with high reliability for any new, unseen data that generates insights for evaluation of the reputation construct based on collected reviews for any (e.g. retail) organisation that is willing to analyse or improve its performance. First, determinant attributes have to be defined in order to generate insights for evaluation with respect to corporate reputation. Second, in order to generate insights data has to be collected and therefore a method has to be developed in order to extract online stakeholder data from reviews. Furthermore, a suitable algorithm has to be created to assess the extracted information based on the determinant attributes in order to analyse the data. Preliminary results indicate that application of our processing pipeline to online employee review data that are publicly available on the web is valid.

Keywords Corporate reputation · Online reviews · Aspect based sentiment analysis (ABSA) · Web scraping

1 Introduction

Reputation has become an important risk concern for many companies worldwide in recent years according to Vig and Dumičić [1]. Reputation represents and creates value for a business. Furthermore, the concern for organisations is that reputation has a sensitive nature and it can take years to build reputation, but it can be destroyed in seconds. A global survey held by Deloitte [2] highlighted reputation risk as top strategic business risk in 2014. This research presented that business executives are finding a loss of reputation one of the biggest risks. Furthermore, a survey conducted by AON Global Risk Management 2019 is underpinning this fact and highlighted that damage to reputation is in the second place, after the risk economic slowdown/slow

R. E. Loke (✉) · IJ. A. A. Steentjes
Centre for Market Insights, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands
e-mail: r.e.loke@hva.nl

recovery [3]. In addition, according to a study conducted by the World Economic Forum, several lines of evidence suggest that reputation accounts for more than 25 percent of the market value of a company [4].

The relevancy of this paper, investigating how corporate reputation can be measured by using online reviews, is twofold. First, according to Resnick [5], a decrease in confidence by investors, employees, clients and other stakeholders are potentially damaging for the sustainability of an organisation in the long-term. An organisation's corporate reputation is the reflection of how it is perceived by its stakeholders. Second, according to Baka [6], online reviews are electronic word-of-mouth and are essential for the reputation management of organisations. Electronic word-of-mouth (reviews) is essential because clients, employees, investors and other stakeholders have become sophisticated users and search online for unbiased information that will help and guide them in their decision-making process. According to Ye and Law [7] and Gruen and Osmonbekov [8], purchase intention and stakeholder's loyalty will be strongly influenced by online reviews. Therefore, a proper corporate reputation management system of online reviews can help an organisation to monitor and analyse its reputation. Ultimately, a company can improve its reputation by using the model which could possibly influence its business performance, positively.

The structure of the paper is as follows. First, we discuss relevant literature in Sect. 2. Thereafter, in Sect. 3, we outline our proposed methodology for measuring corporate reputation and, in Sect. 4, give results that have been obtained for online employee reviews from the public web. We conclude with recommendations in Sect. 5.

2 Literature Review

This literature review gives a relevant overview of recent content of internationally published scientific literature. As already has been outlined in the introduction, reputation is an important risk concern for businesses and CEOs. Furthermore, according to Rindova and Williamson [9] and Dijkmans and Kerkhof [10], reputation is a fundamental intangible asset for companies. The review will mainly focus on the customer and employee-based reputation which are known to be two relevant stakeholders in corporate reputation research. The data for such stakeholders can be publicly collected on customer and employee-oriented platforms on the public web.

2.1 *Definition of Corporate Reputation*

To continue, what does the general term corporate reputation mean? And how can we define it? Several studies have attempted to define the term corporate reputation, but currently there is no general accepted meaning or definition. Whilst the meaning

and definition of the term corporate reputation is arguable, some authors proposed a definition of corporate reputation. A considerable amount of literature has been published on corporate reputation that recognises corporate reputation as a multidimensional construct. However, despite the extensive focus on the subject of corporate reputation in the literature, there is still a lack of consensus on the definition [11]. Table 1 presents an overview of the different definitions used in the literature.

It can be concluded from Table 1 that the existing literature on corporate reputation shows a broad spectrum of definitions. The variation in definitions seems to arise from the different stakeholders that are involved with organisations and the different perspectives seen from varying disciplines.

There is however an overlap between several of these dimensions. One of the aspects of this overlap is that corporate reputation mostly has impact on a few stakeholders or groups of stakeholders. Stakeholder theory suggests, according to Freeman [16], that the expectations towards a company may vary between different stakeholders. Furthermore, “a stakeholder in an organisation is (by definition) any group or individual who can affect or is affected by the achievement of the organisation’s objectives” (2012, pp. 9–10). Using this approach, Wheeler and Sillanpää [17] and Clarkson [18] have been able to distinguish and categorise stakeholders based on their level of interaction with the company and influence in two groups, primary and secondary stakeholders. The former group, according to Vig and Dumičić [1], are interacting with the organisation frequently and include, customers, employees, suppliers, investors and other business partners. The latter does not frequently interact with the organisation and consists of the government, the media, social pressure groups and competitors. The perceived reputation between stakeholders seems to vary, each specific stakeholder seems to experience the corporate reputation in a different way. Therefore, the deeper question could be, does a company has one reputation or many?

A strong perception of the corporate reputation pays off because research has shown that a positive corporate reputation leads to increased sales because it has a positive impact on loyalty and customer purchase intention [19]. Previous research has established according to Lai et al. [20], that brands based on intangible and emotive characteristics are more durable and it seems that competitive erosion has less effect on these brands. Thus, intangible and emotive aspects of a company seem important sources of sustainable competitive advantage. To elaborate on the emotive and intangible aspects that influence the corporate reputation, several lines of evidence suggest that the quality of products, with satisfying service levels that fits the need of a customer, are important in building reputation [21]. Furthermore, an important variable that influences the corporate reputation is the thoughts employees have about their employer [22, 23]. In addition to these variables, according to Fombrun and Ponzi [24], products/services, innovation, workplace, governance, citizenship, leadership, and performance are dimensions to assess reputation. The dimensions that Fombrun mentions [24] have an overlap with the dimensions that Cravens and Oliver [21] and Gatewood and Gowan [22] mention and complement them. These 7 dimensions have been elaborated on in 23 attributes in Fig. 1.

Table 1 Definitions of corporate reputation in the literature

Literature	Definition
Gotsi and Wilson [12], p. 29	“A corporate reputation is a stakeholder’s overall evaluation of a company over time. This evaluation is based on the stakeholder’s direct experiences with the company, any form of communication and symbolism that provides information about the firm’s actions and/or a comparison with the actions of other leading rivals”
Walker [13], p. 370	“A relatively stable, issue-specific aggregate perceptual representation of a company’s past actions and future prospects compared against some standard”
McKenna and Olegario [14], p. 2	“A collective assessment of a company’s attractiveness to a specific group of stakeholders relative to a reference group of companies with which the company competes for resources”
Abimbola [11]	“A stakeholder’s overall evaluation of an organisation over time. This evaluation is based on the stakeholder’s experiences with the organisation and its brand(s), relationships with these and the organisation’s employees and representatives, memberships of brand communities and, any other perceived communication and symbolism that provides information about the organisation’s actions and /or a comparison with the organisation’s rivals”
Fombrun and Gardberg [15]	<p>“Economics Reputations are traits or signals that describe a company’s probable behaviour in a particular situation”</p> <p>“Strategy Reputations are intangible assets that are difficult for rivals to imitate, acquire, or substitute, and so create mobility barriers that provide their owners with a sustained competitive advantage”</p> <p>“Accounting Reputation is one of many types of intangible assets that are difficult to measure but create value for companies”</p> <p>“Marketing Reputation describes the corporate associations that individuals establish with the company name”</p> <p>“Communications Reputations are corporate traits that develop from relationships companies establish with their multiple constituents”</p> <p>“Organisation theory Reputations are cognitive representations of companies that develop as stakeholders make sense of corporate activities”</p> <p>“Sociology Reputational rankings are social constructions emanating from the relationships firms establish with stakeholders in their shared institutional environment”</p>

In the study that Fombrun and Ponzi [24] conducted to measure corporate reputation, Fombrun made a schematic overview of the relation between Corporate Reputation and his RepTrak Pulse model. RepTrak Pulse summarises the seven dimensions into four attributes which are Good Overall Reputation, Good Feeling About, Trust, and Admire and Respect.



Fig. 1 Establishing constructs of corporate reputation: dimensions and attributes

The seven dimensions itself are the result of interviews with reputation managers, senior communication managers of global companies around the world (2015). These executive interviews were supplemented by consumer focus groups that were run in the United States, Europe, and Asia, and nine countries of Latin America [25]. These interviews and focus groups, conducted between 1999 and 2006, were used to identify the seven dimensions of corporate reputation.

Again, these seven dimensions are products/services, innovation, workplace, governance, citizenship, leadership, and performance. Each of these dimensions consist of several underlying attributes. The dimension products/services include whether an organisation offers high-quality products and services that are good value for money. Standing behind its product or services and meeting the customer needs is of importance as well. Together these four attributes represent the dimension products/services. According to Fombrun and Ponzi [24], innovative companies that adapt quickly to change and that launch new products and develop new ideas are more likely to earn respect and admiration and get publications for example in magazines or are being published in rankings of innovative companies. Publications and rankings are visible to stakeholders and thereby adding to an organisation’s reputation. Fombrun’s [24] qualitative research suggests that stakeholders like and respect companies that maintain proper workplaces. Several lines of evidence suggest that satisfied employees are more often committed to long-term involvement and are less likely to turn over. Maintaining a good workplace is critical in recruiting quality workforce for an organisation [26]. The attributes of whether an organisation is

rewarding employees fairly, is offering equal opportunity in the workplace and is demonstrating the concern of health and well-being of its employees are representative for the workplace dimension. Furthermore, governance is determined whether an organisation is perceived as being ethical and transparent. If this is the case, then it's more likely to generate admire and trust for the stakeholders and it will build reputation [27]. With respect to the dimension citizenship, stakeholders tend to respect and admire organisations for their good deeds [28]. In addition, by acting responsibly and by communicating about it, organisations recognise that they are good citizens, and thereby the organisation builds trust and reputation. CEOs can have an important influence for generating admire and trust with the stakeholders [29]. Appealing leaders seem to attract more favourable media coverage and investor endorsements, thereby communicating credibility to stakeholders about company activities and increasing trust in the company, and, thus, building reputation. Performance with respect to the profitability of the company and strong future prospects for growth are important attributes for the performance dimension. Profitability and growth prospect of an organisation have been consistent correlates of reputation in research [30].

Fombrun's [24] RepTrak Pulse model consists of four attributes that have been depicted to the right in Fig. 1: Good Overall Reputation, Good Feeling About, Trust, and Admire and Respect. Analysis of the validity of the four attributes and their ability to assess the overall reputation of an organisation can be incorporated with other measures of interest. The RepTrak pulse model can also be standardised cross-culturally, it can be used with both traditional and online polling techniques, and it can be used to distinguish between corporate reputation and its attributes [31].

These two sides of the RepTrak model correlate with each other.

To underpin again the relevancy for organisations to measure corporate reputation, good reputation management by an organisation that can be assessed for instance from online reviews will result in increased purchase intentions and customer loyalty [32].

2.2 Attributes to Assess Corporate Reputation

As elaborated on in the last section, dimensions to assess corporate reputation could be according to research: the quality of products and services, innovation, workplace, governance, citizenship, leadership and performance as well as the underlying attributes of these dimensions.

In addition, the RepTrak Pulse model [31] defined the attributes Good Overall Reputation, Good Feeling About, Trust, and Admire and Respect to measure corporate reputation.

To extract such attributes from collected review data aspect-based sentiment analysis will be proposed in the next section.

Both sides of Fombrun's model in Fig. 1 will be explored and used to assess and measure corporate reputation.

In the next subsection, the difference between the modelling constructs with respect to the relevancy of the attributes on both sides will be elaborated on by formulating a hypothesis with which the validity of Fombrun's model on online review data can be tested.

2.3 Hypothesis

To be able to test assessing and measuring corporate reputation from online review data we would like to validate whether the two sides that have been mentioned in the conceptual model of the study that Fombrun conducted [24] provide the same score. Therefore, we test the following hypothesis:

H0: A company's corporate reputation measured on online review data with Fombrun's RepTrak Pulse model [24] provides the same score with respect to corporate reputation as the seven dimensions with its underlying attributes.

2.4 Conceptual Model

The conceptual model that we use is based on the attributes defined by Fombrun and the RepTrak model [24] elaborated on in Fig. 1.

The conceptual model consists of two sides: on the left side, corporate reputation and its corresponding seven dimensions and underlying attributes and, on the right side, the RepTrak Pulse model consisting of the four attributes.

In Fombrun and Ponzi [24], these two sides are being presented as two different measurement tools to assess and measure corporate reputation that complement each other with a different intention on different attributes.

The RepTrak Pulse model can be used by practitioners in a lightweight manner to easily assess the status of a firm as well as competitors or business partners.

Assessing lightweight measures of corporate reputation allows organisations to implement market research findings or to leverage reputational advantages [31].

3 Methodology

To be able to assess corporate reputation for organisations, publicly available datasets from an employee-oriented platform at Glassdoor.co.uk have been scraped from the web. An advantage of this platform is that we were able to collect data from multiple companies, like, for example, Albert Heijn, Bol.com, Booking.com, and Philips. A limitation stems from the fact that not all organisations post their reviews publicly online on this platform; also, there are organisations that never, or barely, post any

reviews online. However, it can be argued that stakeholder transparency will be beneficial for organisations in the long run.

The review datasets that were collected have been assessed with unsupervised Aspect Based Sentiment Analysis based on Bidirectional LSTM (hereafter BiLSTM) [33]. Using an unsupervised ABSA model increases the efficiency because there is no need to label the reviews into a specific category. The ABSA model used in our work has been developed by Intel AI. [34] and is using state-of-the-art topologies (e.g. Spacy, TensorFlow, and PyTorch) to extract corresponding aspects and opinions from online reviews. The authors of this paper had contact with the architects of this model to adjust and tune the model according to the needs of our project. Intel's ABSA model seems impossible to surpass in terms of quality, especially within the given timeframe of small projects.

In the following subsections, we describe the data collection and data processing that is involved in our corporate reputation measurement pipeline.

3.1 Data Collection

The primary data used in our project is collected from employee-oriented platform (Glassdoor.co.uk). Data has been collected from English and Dutch-speaking employees. There hasn't been a distinction made between cross-cultural differences, and according to the findings in [31] with respect to the RepTrak Pulse model, it seems that this isn't necessary while comparing and analysing the data.

However, not all the data that we collected was used for analysis, only the collected Booking.com dataset has been used. Due to limited time and the time it takes to analyse a dataset with the proposed model (takes ~6 h to run one dataset) one dataset has been analysed to give an indication about the model's performance.

The dataset that has been used for the analysis consists of 2391 reviews written in English. This dataset has been used because it has the most reviews, is the most complete dataset from the companies that have been scraped and therefore the biggest sample set to test our hypothesis. A larger sample size of reviews is affecting the accuracy and reliability of the outcomes in a positive way. Therefore, the dataset of Booking.com has been used to get the best results possible from using reviews with the proposed model.

The data has been retrieved by using a web scraping tool [35] based on a .NET framework to systematically mine and collect data to limit the chance of human error [36]. It can extract images, text, and files from web pages and the data can be exported into multiple formats like .CSV (Comma Separated Value), Microsoft Access, and in an SQL Database.

This scraping tool has been used due to its ease-of-use and the fact that developing and building a web scraper in Python takes time. EasyWebExtract takes less time to deploy than a customised Python solution and the time saved is used to be able to setup and apply the Aspect Based Sentiment Analysis (ABSA) method.

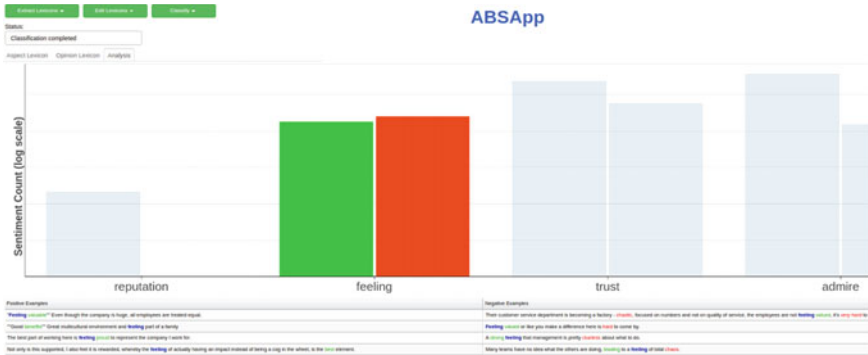


Fig. 2 Example of reviews RepTrak side

A limitation with respect to the results is the lack of spread in the data. Only one dataset and only one company have been used to test our hypothesis.

Figure 2 shows several reviews from the scraped dataset as well as relevant processing on the RepTrak side that we will explain in the next subsection.

3.2 Data Processing

The review data was analysed with aspect-based sentiment analysis (hereafter ABSA) method that has been used to derive an in-depth understanding of the collected reviews and to extract meaningful aspects from the reviews.

According to Intel [34], ABSA is the process of co-extracting opinion terms and aspects, and their relation within a given corpus. For our purpose, these aspects should preferably be similar to the variables and attributes from Fombrun and Ponzi [24] RepTrak model as well as his corporate reputation model.

During the quest of finding an applicable processing method, we found, after extensive search and consideration, the Python library NLP Architect [34]. This ABSA method is unsupervised, which means to some extent that the user doesn't have to label the data, which is labour intensive, especially when analysing thousands of reviews [37]. A huge advantage of unsupervised processing is that you are able to analyse raw review data right away.

The process flow in the ABSA method is as depicted in Fig. 3.

The first training step is pre-processing of the review text and is performed by a package called spaCy [38]. During this step the text is being tokenized meaning that a paragraph or review is splitted into sentences and sentences into words, a word is then a token in the sentence.

The second part is applying a dependency parser to the training data. A dependency parser establishes grammatical relationships within the structure of a sentence [39]. The NLP Architect package makes use of a parser that's based on BiLSTM [33].

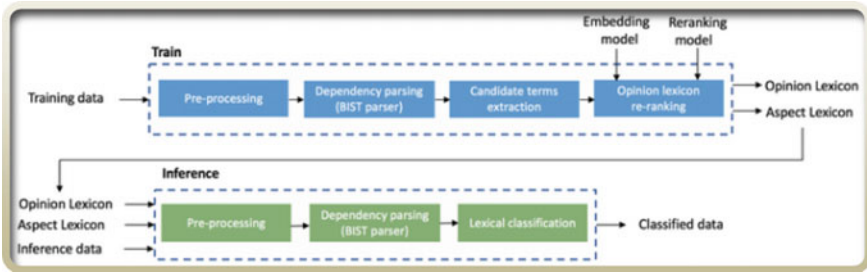


Fig. 3 ABSA processing pipeline, overview

The third part is applying a lexicon acquisition algorithm on the training data that is based on double propagation as described in [40]. An advantage of double propagation is that it only needs the initial lexicon acquisition to start the process. According to Qiu and Liu [40], this method outperforms several compared existing methods significantly. Furthermore, the algorithm uses a generic lexicon introduced in [41].

The last step includes a Multilayer Perceptron and re-ranks the opinion term and polarity estimation algorithm.

The inference phase inputs the selected dataset together with the opinion and aspect lexicon. The output of this phase is a list of aspect-opinion pairs complemented with polarity and score. The inference phase is used to detect related aspect-opinion pairs.

Figure 4 shows the aspect lexicon that we applied on the RepTrak side.

Extract Lexicons ▾
Edit Lexicons ▾
Classify ▾

Status:

Classification completed

Aspect Lexicon

Opinion Lexicon

Analysis

#	<input checked="" type="checkbox"/>	Term	Alias1	Alias2	Alias3
0	<input checked="" type="checkbox"/>	reputation			
1	<input checked="" type="checkbox"/>	feeling			
2	<input checked="" type="checkbox"/>	trust			
3	<input checked="" type="checkbox"/>	admire	respect		

Fig. 4 Aspect lexicon, RepTrak side

4 Results

The proposed hypothesis has been tested on the collected Booking.com dataset. The reason for using only the Booking.com dataset has been elaborated on in Sect. 3.1.

The results on the right side of Fombrun’s model [24] are depicted in Fig. 5.

The amount of positive and negative reviews on the attributes of the RepTrak side are depicted in Table 2.

The results on the left side of Fombrun’s model [24] are depicted in Fig. 6.

The amount of positive and negative reviews on the attributes of the corporate reputation side are depicted in Table 3. Not all the attributes of Fombrun’s RepTrak model [24] are present in the reviews within the dataset, therefore not all attributes are present in this table.

The equal weight method [42] has been used to determine the relation between the two sides. The calculation to determine the differences in results has been established by dividing the positive reviews by the sum of positive and negative reviews with respect to a specific attribute. This resulted in a difference of 7.42% in outcome

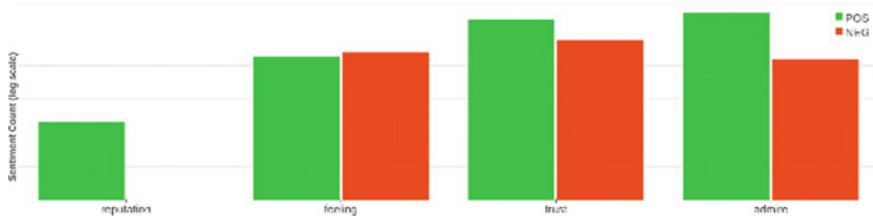


Fig. 5 RepTrak side score on the attributes for booking.com

Table 2 Four attributes side

Attribute	Reputation	Feeling	Trust	Admire and respect	Score
Positive	5	19	41	47	
Negative	0	21	27	18	
Score 0–1	1.00	0.48	0.60	0.72	0.70

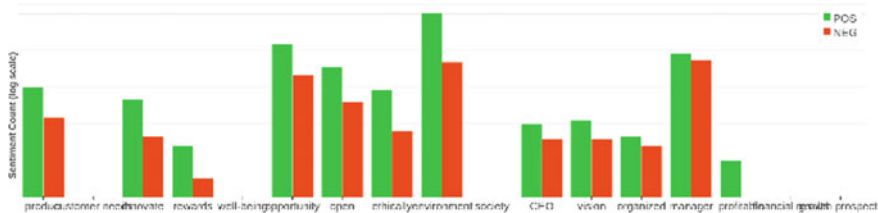


Fig. 6 Results booking.com Corporate Reputation side

Table 3 Seven dimensions side

Dimension	Attribute	Positive	Negative	Score 0–1
Products	Product	62	20	0.76
Innovation	Innovate	39	10	0.80
Workplace	Rewards	7	2	0.78
Workplace	Opportunity	319	98	0.76
Governance	Open	134	36	0.79
Governance	Ethically	56	12	0.82
Citizenship	Environment	1026	161	0.86
Leadership	CEO	16	9	0.64
Leadership	Vision	18	9	0.67
Leadership	Organized	10	7	0.59
Leadership	Manager	225	176	0.56
Performance	Profitable	4	0	1.00
Overall				0.75

between the two sides of Fombrun’s model [24]. This outcome could be considered to reflect a positive correlation.

5 Recommendations

In the following paragraphs, we give some recommendations that are related to our work.

We advise that decision makers in companies measure and assess corporate reputation by using the mentioned ABSA method to extract the seven dimensions and the four attributes that are mentioned in the RepTrak Pulse model [24] from online reviews.

Assessment of the seven dimensions from the RepTrak Pulse model [24] will, in general, agree in outcome with assessment of the four attributes from the RepTrak Pulse model. Both assessments can bring to light various relevant signals to companies. Interpretation of the signals should be carefully guided in order to be able to steer and influence decision making.

Our processing pipeline and the ASBA method employed can be used to determine corporate reputation but also wider in society which is of vital importance since the amount of online reviews is increasing sharply and has become more important for decision makers [6].

Reputation accounts for more than 25 percent of the market value of a company [4]. Therefore, measuring corporate reputation has become important for organisations and possibly can become of vital importance for a company’s existence. Ultimately,

based on established corporate reputation (measurements), a company could even improve its reputation.

The seven dimensions and the four attributes that we use in our approach are all relevant to corporate reputation. A company can assess its reputation for each attribute or determinant.

A recommendation for further research is to test the proposed hypothesis on multiple companies and datasets to strengthen the reliability.

Similarly, another recommendation for further research is to assess and validate the mutual correlation between all the attributes both at the left and at the right of Fombrun's model [24] to demonstrate the reliability of online data.

Corporate reputation has been tested on employee stakeholder data, if tested on other stakeholder data, customers, for example, results could differ.

Similarly, the ABSA model has been tested in this study based on online reviews from employees, further research is needed to test whether this model can also be applied to other stakeholder reviews to measure corporate reputation.

More research is needed with respect to fake online employee reviews and how they affect the outcome of the measurements.

Acknowledgements This paper has been inspired on the MSc master project of IJsbrand Steentjes who was involved via the master Digital Driven Business at HvA. Thanks go to Jesse Weltevreden for providing some useful suggestions to an initial version of this manuscript. Rob Loke is assistant professor data science at CMIHvA.

References

1. Vig, S., Dumičić, K.: The impact of reputation on corporate financial performance: median regression approach. *Bus. Syst. Res.* **8**(2), 40–58 (2017)
2. Deloitte: Retrieved from 2014 global survey on reputation risk: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Governance-Risk-Compliance/gx_gr_Reputation@Risk%20survey%20report_FINAL.pdf (2014)
3. AON: Global Risk Management Survey. Retrieved from AON: <https://www.aon.com/2019-top-global-risks-management-economics-geopolitics-brand-damage-insights/index.html> (2019)
4. World Economic Forum: Global Risks 2012, 7th edn. Retrieved from http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2012.pdf(2012)
5. Resnick, J.T.: Managing corporate reputation—Applying rigorous measures to a key asset. *J. Bus. Strategy* **26**(6), 30–38 (2004)
6. Baka, V.: The becoming of user-generated reviews: looking at the past to understand the future of managing reputation in the travel sector, vol. 53, pp. 148–162. Elsevier (2016)
7. Ye, Q., Law, R.: The influence of user-generated content on traveller behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Hum. Behavior* **27**(2), 634–639 (Elsevier).
8. Gruen, T.W., Osmonbekov, T.: eWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty. *J. Bus. Res.* **59**(4), 449–456 (2006) (Elsevier)
9. Rindova, V.P., Williamson, I.O.: Reputation as an intangible asset: reflections on theory and methods in two empirical studies of business school reputations. *J. Manage.* (2010)
10. Dijkmans, C., Kerkhof, P. K.: A stage to engage: social media use and corporate reputation, vol. 47, pp. 58–67. Elsevier (2015)

11. Abimbola, T.T.: Corporate identity, corporate branding and corporate reputations: reconciliation and integration. *Eur. J. Mark.* **46**(7–8), 1048–1063 (2012).
12. Gotsi, M., Wilson, A.M.: Corporate reputation: seeking a definition. *Corp. Commun.* **29** (2001)
13. Walker, K.: A systematic review of the corporate reputation literature: definition, measurement, and theory. *Corp. Reput. Rev.* **370** (2010)
14. McKenna, C., Olegario, R.: Corporate reputation in historical perspective. *Bus. History Rev.* (2012)
15. Fombrun, C.J., Gardberg, N.A.: The reputation QuotientSM: a multi-stakeholder measure of corporate reputation. *J. Brand Manage.* **7**, 1–15 (2000)
16. Freeman.: Strategic management: a stakeholder approach. Pitman Publishing Inc., Boston (1984)
17. Wheeler, D., Sillanpää, M.: *The Stakeholder Corporation: A Blueprint for Maximizing Stakeholder Value*, 1st edn. Ft Press, Upper Saddle River, New Jersey, USA (1997)
18. Clarkson, M.: A stakeholder framework for analyzing and evaluating corporate social performance. *Acad. Manage. Rev.* (1995)
19. Caruana, A., Ewing, M.T.: How corporate reputation, quality, and value influence online loyalty. *J. Bus. Res.* **63**(9–10), 1103–1110 (2010)
20. Lai, C.S., Chiu, C.J., Yang, C.F.: The effects of corporate social responsibility on brand performance: the mediating effect of industrial brand equity and corporate reputation. *J. Bus. Ethics* (2010)
21. Cravens, K., Oliver, E.: The reputation index: measuring and managing corporate reputation. *Eur. Manage. J.* **21**(2), 201–212 (2003)
22. Gatewood, R.D., Gowan, M.A.: Corporate image, recruitment image, and initial job choice decisions. *Acad. Manage. J.* **36**(2), 414–427 (1993).
23. Helm, S.: Employees' awareness of their impact on corporate reputation. *J. Bus. Res.* **64**(7), 657–663 (2011)
24. Fombrun, C.J., Ponzi, L.J.: Stakeholder tracking and analysis: the RepTrak[®] system for measuring corporate reputation. *Corp. Reput. Rev.* **18**(3–24), 3–24 (2015)
25. Alloza, Á., Carreras, E.: *Corporate Reputation*. LID Publishing, London (2013)
26. Alniacik, E., Alniacik, U.: How do the dimensions of corporate reputation affect employment intentions? *Corp. Reput. Rev.* **15**(1), 3–19. Springer
27. Soleimani, A., Schnepfer, W.D.: The impact of stakeholder power on corporate reputation: a cross-country corporate governance perspective. *Org. Sci.* **25**(4), 991–1008 (2014)
28. Orlitzky, M., Swanson, D.: Assessing stakeholder satisfaction: toward a supplemental measure of corporate social performance as reputation. *Corp. Reput. Rev.* **15**(2), 119–137 (2012).
29. Halff, G.: The presentation of CEOs in economic downturn. *Corp. Reput. Rev.* **16**(3), 234–243 (2013)
30. Lange, D., Lee, P.: Organizational reputation: a review. *J. Manage.* **37**(1), 153–185 (2011)
31. Ponzi, L.J., Fombrun, C.: RepTrakTM pulse: conceptualizing and validating a short-form measure of corporate reputation. *Corp. Reput. Rev.* **14**(15–35), 15–35 (2011) (Springer)
32. Liu, Z. Park, L.: What makes a useful online review? Implication for travel product websites. *Tour. Manag.* **47**(1), 140–151 (2015)
33. Goldberg, E.K.: Simple and Accurate Dependency Parsing. Association for Computational Linguistics, pp. 313–327 (2016)
34. Intel: NLP Architect by Intel[®] AI Lab. Retrieved from nlp_architect.nervanasys (2020)
35. EasyWebExtract: What is Easy Web Extract? Retrieved from EasyWebExtract: <http://webextract.net/> (2020)
36. Vargiu, E., Urru, M.: Exploiting web scraping in a collaborative filtering- based approach to web advertising. *J. Artif. Intell. Res.* **2**(1), 44–54 (2013)
37. Erhan, D., Bengio, Y.: Why does unsupervised pre-training help deep learning? *J. Mach. Learn. Res.* 625–660 (2010)
38. spaCy: spaCy. Retrieved from Industrial-Strength Natural Language Processing: <https://spacy.io/> (2020)

39. Stanford: The Stanford NLP Group. Retrieved from Stanford: <https://nlp.stanford.edu/software/nndep.html> (2020)
40. Qiu, G., Liu, B.: Opinion word expansion and target extraction through double propagation. *Comput. Linguist.* 9–27 (2011)
41. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: *Conference on knowledge discovery and data mining*, pp. 168–177 (2004)
42. Jia, J., Fischer, G.: Attribute weighting methods and decision quality in the presence of response error: a simulation study. *J. Behav. Decis. Mak.* **11**(2), 85–105 (1998)

Online Reviews of Discount Products: The Case of Steam



Bruno Sista, Beatriz Casais , and Nuno Moutinho 

Abstract This paper analyzes the role of price in the consumers' pre-purchase expectations and post-consumption evaluations. The researchers observed, in the Steam platform, how promotional discounts can lead to fluctuations in user recommendations for discount products. Applying the change point analysis method for the observation of longitudinal data, the results indicate that temporary discounts can disrupt the stable process of word of mouth generation. There is a significant effect on the volume of reviews posted for a product shortly after the occurrence of a discount and also variations in review scores that can be either positive or negative.

Keywords Price promotions · Electronic word of mouth · User reviews · Steam · Online rating · Price discounts

1 Introduction

The advent of the commercial Internet allowed for the development of new communication tools and several new forms of interaction between users online. The Web 2.0, in turn, has redefined the Internet, giving its users new powers to contribute with content to the network and bringing a social component to most online experiences [1]. Several services have appeared on the Web appealing to the “participative culture,” from blogs to social networks, inviting the user to take part in the creation of social, economic and cultural value. e-Commerce is perhaps one of the most impactful new paradigms to emerge from this trend. The appearance of dotcoms such as Amazon.com and ebay.com, pioneers in the creation of online marketplaces that facilitate the fulfillment of transactions through the Web has come to revolutionize the way citizens of the Internet (individuals or companies) relate to trade and

B. Sista · N. Moutinho
Faculty of Economics and CEFUP, University of Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

B. Casais (✉)
School of Economics and Management and CICS.NOVA.UMinho, University of Minho, Campus de Gualtar, 4700-057 Braga, Portugal
e-mail: bcasais@eeg.uminho.pt

shopping. The performance of these marketplaces is dependent on the participation of consumers, their main selling point being the direct interaction between sellers, prospects and buyers, which sets them apart from physical points of sale [2]. Social Commerce is a relatively recent concept, appearing under the umbrella of the Web 2.0 that represents business made through social media [2]. The modern user is no longer limited to the information producers/retailers choose to make available for their merchandise, or articles published in the specialized press, instead sourcing his information to his peers, who share first-hand opinions and reviews on their experiences with those products.

The creation of this collective intelligence has brought such impact on the decision-making processes for online shopping, that electronic word of mouth (henceforth, eWOM) and user-generated Content (UGC) have become two of the most discussed subjects between marketers, who now more than ever recognize the effects of customers participating in the creation of brand equity, either by increasing notoriety and reach of the brand, or through the part they play in changing other users' perception of quality of the products [3]. Customer feedback management and monitoring of user reviews in online platforms must be a concern addressed in marketing since these public manifestations of word of mouth can have very a positive effect on the generation of leads and business, but also a notably negative impact in the event negative buzz starts to circulate online.

Although recent literature recognizes consumer-to-consumer (C2C) communication as having a relevant impact on the purchasing decision-making of e-Commerce platform users, with the potential to affect volumes of sales [4], there is still a lack of research on how word of mouth reacts to mutations in the marketing mix. The aim of this article will be to analyze how pricing changes (specifically through promotional discounts) can lead to fluctuations in the user review scores of products in an online marketplace. This will hopefully shed some light on the role of price in the customers' pre-purchase expectations, and the post-purchase evaluation of their consumption experience.

2 Literature Review

Social commerce is a recent concept, appearing under the umbrella of the Web 2.0 [2, 5]. The modern user is no longer limited to the information producers/retailers choose to make available for their merchandise and services, articles published in the specialized press, but they tend to source for information from peers, who share first-hand opinions and reviews of their experiences with those products [3, 6]. The creation of this collective intelligence has brought such impact on the decision-making processes for online shopping, that electronic word of mouth (eWOM) and user-generated content (UGC) have become two of the most discussed subjects between marketers [6, 7], who now more than ever recognize the effects of customers participation in the creation of brand equity, either by increasing notoriety and reach of the brand,

or through the part they play in changing other users' perception of quality of the products.

Recent literature recognizes C2C communication has a relevant impact on the purchasing decision-making of e-Commerce platform users, with the potential to affect volumes of sales [4]. Since that pioneer work, several studies have applied the same principles to different markets, the main consensus being that online shoppers adopt online reviews as reliable information sources in their decision-making [5, 7–9].

Online marketplaces have recognized the potential of having recommendations systems and public review systems, where consumer opinions are gathered and often condensed in the form of review scores, or ratings, giving the customer a greater range of information about the products on sale without having to visit other websites [8, 9].

Besides users perceiving value in verbalizing their opinions online for the same motives we have discussed earlier, users read those opinions. The most common goals are to save decision-making-time and make better decisions, but there are other factors such as remuneration or the sense of belonging to a community that drive users to consume word of mouth [6, 7].

Furthermore, the fact that communication in online platforms is durable and information is stored and visible for most visitors to read, brings a new light to the community involvement motivation for engaging in eWOM [9]. Here, more than in traditional discussions that contribution is not completely altruistic. By sharing their opinions through the Internet, users of Web-based opinion platforms hope to motivate others to do the same, increasing the flux of information. That way, for example, if they encounter some problem with their product, they are more likely to find someone who has encountered the same issue before and knows how to solve it. And, in turn, that is more likely to happen once the platform has reached a critical mass of users and interactions [6, 8].

The idea that online reviews influence the consumer's decision-making process has been proven true by several studies. Online recommendations acted as reliable information sources in the pre-purchase information seeking phase. Online word of mouth has been suggested by several authors as a promising variable in forecasting future sales [10]. One particular study, regarding daily box office performance for recently released movies, concluded that average online review ratings had no influence in future daily revenues, suggesting that the scores attributed by other users had no persuasive effect over consumers' decisions of which movie(s) to watch [11]. However, the same study confirmed that the increased volume of online posting alone had a positive impact on revenues due to the awareness effect, as online reviews are an indicator of intensity of word of mouth, which is a significant driver for movie sales. Chevalier and Mayzli [4] deserve a special mention for their pioneer analysis of user reviewing behavior through the observation of aggregate review scores and sales figures across two different online book stores (Amazon.com and Barnes and Noble's website bn.com) over the course of a year. They concluded that changes in user reviews had a significant effect on sales at least on Amazon.com, not only if you consider fluctuations in average review scores, but also based on the percentage of negative opinions.

Users who share their opinion in search of recognition, self-enhancement or social benefit, can also see the Internet as a facilitator. Online marketplaces often recognize a user's frequency of activity through reputation systems, letting shoppers access other users' review history, which ultimately allows the perception of expertise and how helpful they have been in the past. There is a competition for attention between online reviewers and their strategic perspective on content creation.

Despite the impact of eWoM in the success of e-Commerce strategies, and the fact that price is generally accepted as a relevant determinant of customer satisfaction; we found a lack of research on the impact of pricing on online product reviews. Li and Hitt [12] confirmed a correlation between price and product evaluations for cross-sectional datasets.

In steam, a digital games distribution platform, consumer reviews are made on a two-option scale: users either recommend a product they purchase, or recommend against it. The outcome is then aggregated and translated into a 9-point scale, ranging from "Overwhelmingly Negative" to "Overwhelmingly Positive" based on the total number of reviews and the percentage of reviews that are positive. This percentage is also shown on the platform's product search engine and can be used as a sorting criterion to order products from the most recommended to least recommended. The absolute number of positive and negative reviews can also be seen on the product page.

The steam reviewing system invokes the notion of a certain standard of quality that is pre-formed by consumers based on price and quality perceptions and that generate feelings of satisfaction or discomfort when compared to the actual perceived quality of products post-consumption. To understand this mechanism, it is particularly useful to analyze the thesis by Voss et al. [10] which, although primarily focused on service exchanges, is among the first articles to emphasize the role of pricing in predicting customer satisfaction.

3 Methodology

The goal of this research is to confirm a relationship of causality between pricing initiatives activated on a given product (specifically, promotional discounts) and the generation of electronic word of mouth (through the proxy of online product reviews) for that product, analyzing the behavior of user reviews after a variation in price.

This exploratory study was translated into the following hypotheses:

F1 Hypothesis 1 The occurrence of a price reduction in a product lead to an increase in the number of user reviews published in the days following the discount.

F1 Hypothesis 2 The occurrence of a price reduction causes a change in the average review score of a product in the days following a discount.

To test these hypotheses, our study focuses on the behavior of users of the steam platform, a software digital distribution service which has been developing efforts in

making the purchase and consumption of videogames into a social experience. The choice of a digital distribution service is justified because it allows us to eliminate as many third-party intermediaries as possible between producers and consumers. Other Web services such as Amazon or eBay, because they mostly sell physical products, could lead to distortions on user review scores due to users biasing their product reviews in the light of logistical distribution problems (delivery delays, damaged products, stock ruptures, etc.). Steam users are provided with a homogeneous product, and the variations in consumption experience derive only from the different hardware setups which they use to play. Computerized scraping methods were used to download data from the Steam storefront search engine (store.steampowered.com/search), through daily extractions between the January 18 and March 30, 2015. The following variables were stored for over 8.000 products on sale, for each of the 71 daily observations:

- Gross price
- % of promotional discount
- Total no. of user reviews
- % of positive reviews.

These variables were considered to be the most adequate for a first exploration of a relationship between pricing initiatives and word of mouth (volume and nature), mainly for their exposure in the Steam platform, being the most emphasized and readily available proxies of the constructs under analysis.

Although data was collected for all products in the Steam storefront, we opted to select a sample that we assume most closely represents the behavior of the market under regular circumstances. For that purpose, we selected only 658 products that fulfilled a series of criteria of “normality” (e.g., not free, released before 2015, ≥ 100 reviews).

For each of the products, we extracted the 29 observations that spanned the period between 14 days before and 14 days after the occurrence of the discount. In this timeline, and for the interpretation of the graphics and figures below, consider $T = 15$ as the first day of the price discount.

To test the hypotheses, we analyzed the evolution of two main variables:

- **New reviews published daily.** Considering that, we are able to extract the total number of reviews existing at the time of extraction:

R_T Total number of user reviews existing on day T .
 r_T Number of new reviews published on day T

$$R_T = R_{T-1} + r_T \tag{1}$$

- **Average product review score** is defined by the percentage of all reviews posted up until the time of extraction that are positive:

S_T Average review score on day T , i.e.,

P_T Total number of positive reviews existing on day T .

p_T New positive reviews published on day T

$$S_T = \frac{P_T}{R_T} = \frac{P_{T-1} + p_T}{R_{T-1} + r_T} \tag{2}$$

Hence, the data collected should be sufficient to detect periods in which the average score of reviews posted was higher (or lower) than the sum of all reviews posted up until that point in time.

The serially dependent nature of these variables defies the assumption of i.i.d. observations for the change point analysis method that will be used to detect changes in means. This will lead to an increased vulnerability to type I errors because of the possibility of overestimation of residuals [13]. However, this assumption is one that is frequently overlooked by literature in the analysis of empirical data since real-time series are rarely stationary and truly stochastic processes.

4 Results

Figure 1 depicts the evolution of this variable, where it is easy to detect that an outstanding number of reviews are published in the few days after the start of a price

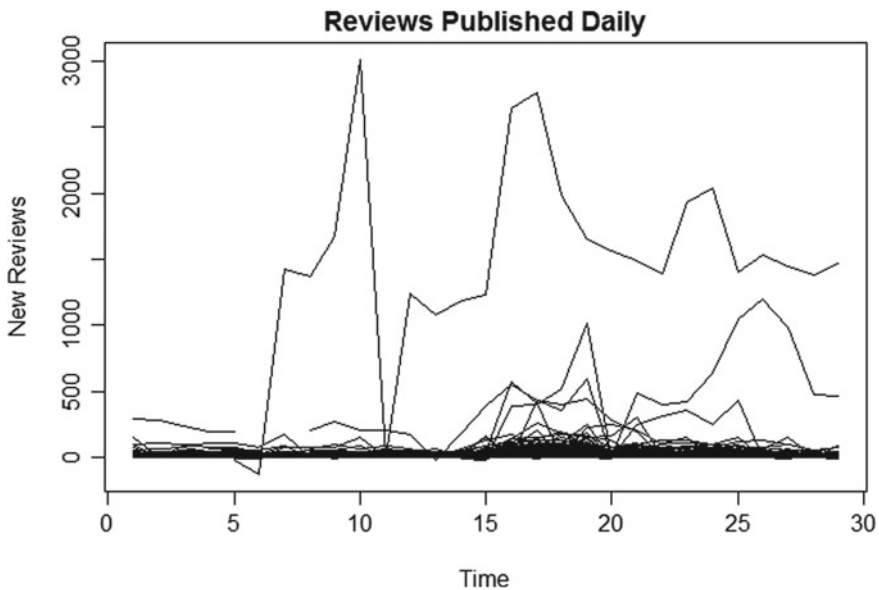


Fig. 1 Number of reviews published daily

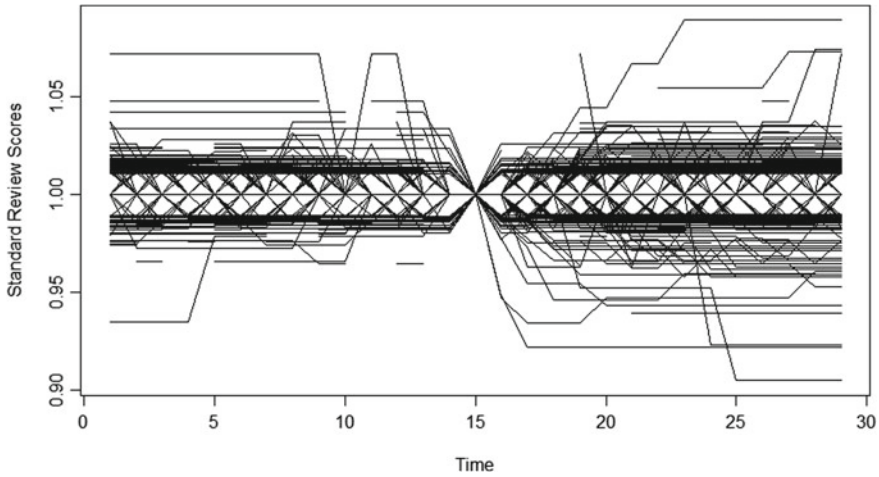


Fig. 2 Standard review scores

discount ($T = 15$), when compared with the relatively stable process of generation of user recommendations. Testing for a change point gives a positive result for a change in means in $T = 16$ with at least 99% confidence. Thus, we confirm Hypothesis 1, and conclude, about the behavior of consumers in the Steam platform, that a price reduction has a significant impact on the volume of product reviews published in the days following the pricing initiative.

For review scores, the result provided some information about the behavior of our variables. One of the most evident interpretations of this plot is that average review scores are relatively stable for the period of analysis, which is coherent with the fact that we are analyzing cumulative, serially dependent scores. Our data is also presented in a percent-point scale, which explains the predominantly parallel lines. This chart was certainly not sufficient to answer our research questions; although it was immediately apparent that review scores are subject to more frequent variations after $T = 15$ (incidentally, the first day of the price discount), as we can observe from the higher density of segments with non-null slope in the second half of the time series plot.

Although review scores range from 12 to 100% for the period observed, they are highly concentrated around the 85% mark and the distribution is relatively stable.

To visually interpret our data, we found it useful to compare the review score of each product at a certain moment in time, S_t , to the review score it had on the moment of occurrence of the discount.

For that matter, we create yet another construct—the Standard Score (S'_T):

$$S'_T = \frac{S_T}{S_{15}} \tag{3}$$

Results in this new time series chart (Fig. 2) seems to confirm that after the occurrence of a discount ($T = 15$), there is a higher volatility in customer reviews scores than before, and that in that case the signal of variation can be either positive or negative.

We recognized a pattern for a considerable number of time series (products) coherent with the existence of a change point. In fact, testing for a change point of the distributional means with the *ecp* package returns positive results for a first-level change point at $T = 17$ with at least 99% confidence, equivalent to dividing the time series into two clusters, $T = [1, 16]$ and $T = [17, 29]$ with statistically significant differences in means. This result is coherent with the previous analysis of the CUSUM control chart, which is noticeably skewed toward the end of the time series (and to the right of $T = 15$, the first observation with presence of a price discount). This can easily be justified because of: (1) the delay between the moment a discount is activated and the moment a consumer finally publishes his review, after experiencing and evaluating the product; (2) the serial dependence of the variable in cause, paired with some rigidity in the scale (percentages with no decimal places) which provides resistance and causes a delay in variations of average review scores.

The confirmation of a statistically significant change point in the time series around the occurrence of a price discount is deemed sufficient to accept Hypothesis 2. Thus, we conclude that, regarding review scores for products in the Steam storefront, the occurrence of a price discount causes a shift in the means of the distribution. Regarding the signal of that shift, analyzing a sub-sample of products that have review information available for both $T = 14$ (last observation before a discount) and $T = 29$:

- 103 products for which $S_{14} > S_{29}$
- 325 products for which $S_{14} = S_{29}$
- 83 products for which $S_{14} < S_{29}$.

5 Conclusions

Applying the change point analysis method for the observation of longitudinal data, we confirm that temporary discounts can disrupt the otherwise stable process of word of mouth generation. They have not only a significant effect on the volume of reviews posted for a product shortly after the occurrence of a discount, but also cause variations in review scores that can be either positive or negative. Regardless of the direction of variation of review scores, it is implied by several studies that online word of mouth and customer feedback are significant factors in the long-term generation of business and can be used to forecast future sales. If a price promotion has a durable effect in the review scores, and future buyers use those scores as quality cues in the information gathering phase of their purchasing process: a temporary price reduction made with the intent to generate a durable raise in online review scores could have a positive return in the long-term; a temporary price reduction made with the intent of generating a short-term increase in sales and/or awareness (by volume of word of

mouth) could generate a decrease in review scores that has a negative impact on the long run and in the price that customers are willing to pay in the future due to the decrease in value perceptions.

Further exploration of how online word of mouth is generated and which factors contribute to an increase or decrease of review scores in the presence of pricing initiatives could lead to developments in the paradigms of e-commerce and social commerce.

6 Limitations and Future Research

This study should be interpreted considering that Steam is a software digital distribution service, a typical supply-side market with a posted-price context. When the seller determines a price reduction, there is an increase in the total consumer surplus leading to an increase in the volume of sales with no decrease in quality—there is an infinite supply of homogeneous product [14]—which is not always the case for retailers of physical goods. As for review scores, our study showed that the effect of price promotions in review scores is not linear, and understanding which factors can influence the signal of variation in user ratings should serve as the motto for future research. Using sentiment analysis methods to identify differences in the contextual cues of product reviews published before and after a price promotion can be key to understanding the consumers' thought processes when making a purchasing decision online. New customers may account for the temporary nature of the reduction when creating their reviews and still evaluate their purchased products based on the gross (non-discounted) price instead of the price at which they acquired them. This would lead to “false-positive”/“false-negative” reviews, which present a conditional evaluation based on price [12]. According to Darke and Chung [15], price promotions have negative effects in the consumers' value perceptions of products. Introducing a price reduction could undermine a customer's appreciation of quality and online feedback. Festinger's dissonance theory [16] implies that consumers tend to raise their evaluations of products when their cost of acquisition is higher.

References

1. Balio, S., Casais, B.: Content marketing framework to analyze customer engagement on social media. In: *Research Anthology on Strategies for Using Social Media as a Service and Tool in Business*, Chapter 16, pp. 320–336. Information Resources Management Association, IGI Global (2021)
2. Terra, L., Casais, B.: Moments of truth in social commerce customer journey: a literature review. In: Martínez-López, F.J., D' Alessandro, S. (eds.). *Advances in Digital Marketing and eCommerce*, DMEC 2021, pp. 236–242. Springer (2021)
3. Oliveira, B., Casais, B.: The importance of user-generated photos in restaurant selection. *J. Hosp. Tour. Technol.* **10**(1), 2–14 (2019)

4. Chevalier, J., Mayzlin, D.: The effect of word of mouth on sales: online book reviews. *J. Mark. Res.* **43**, 345–354 (2006)
5. Hajli, N.: Social commerce constructs and consumer's intention to buy. *Int. J. Inf. Manage.* **35**, 183–191 (2015)
6. Schafer, J.B., Konstan, J., Riedi, J.: Recommender systems in e-commerce. In: *Proceedings of the 1st ACM Conference on Electronic Commerce EC-99*, pp. 158–166. Denver (1999).
7. See-To, E.W.K., Ho, K.K.W.: Value co-creation and purchase intention in social network sites: The role of electronic word-of-mouth and trust—A theoretical analysis. *Comput. Hum. Behav.* **31**, 182–189 (2014)
8. Cui, G., Lui, H., Guo, X.: The effect of online consumer reviews on new product sales. *Int. J. Electron. Commer.* **17**, 39–58 (2012)
9. Filieri, R., McLeay, F.: E-WOM and accommodation: an analysis of the factors that influence travelers' adoption of information from online reviews. *J. Travel Res.* **53**, 44–57 (2014)
10. Voss, G.B., Parasuraman, A., Grewal, D.: The roles of price, performance, and expectations in determining satisfaction in service exchanges. *J. Mark.* **62**, 46–61 (1998)
11. Duan, W., Gu, B., Whinston, A.B.: Do online reviews matter?—An empirical investigation of panel data. *Decis. Support Syst.* **45**(4), 1007–1016 (2008)
12. Li, X., Hitt, L.M.: Price effects in online product reviews: an analytical model and empirical analysis. *MIS Q.* **34**, 809 (2010)
13. Lund, R., Wang, L., Lu, Q.Q., Reeves, J., Gallagher, C., Feng, Y.: Changepoint detection in periodic and autocorrelated time series. *J. Clim.* **20**, 5178–5190 (2007)
14. Kleinberg, R., Leighton, T.: The value of knowing a demand curve: bounds on regret for online posted-price auctions. In: *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*. Cambridge (2003)
15. Darke, P.R., Chung, C.M.: Effects of pricing and promotion on consumer perceptions: it depends on how you frame it. *J. Retail.* **81**, 35–47 (2005)
16. Festinger, L.: A theory of cognitive dissonance. *Sci. Am* 207 (1957).

Cough Sound Identification: An Approach Based on Ensemble Learning



Christian Salamea-Palacios , Javier Guaña-Moya , Tarquino Sanchez ,
Xavier Calderón , and David Naranjo

Abstract Cough identification using DSP techniques in an audio signal is a complex task; thus, an artificial intelligence approach is proposed by applying machine learning, deep learning, and HMMs algorithms. Later, an ensemble learning model has been used to differentiate cough from other environmental sounds, putting those algorithms together and choosing the best result as the performance of the system. The final system consists of a preprocessing stage where the audio signals are adjusted through data augmentation, normalization, removal of silent fragments, and the transformation to Mel spectrograms, while, on back-end stage, three models have been evaluated: a convolutional neural network, a random forest, and a classifier based on hidden Markov models. We assembled a hard voting classifier (VC) model from the three models to obtain a more robust and reliable model. The VC model reached the highest precision and F1-score values without false-negative and up to 75% of true-positive values.

Keywords Cough identification · Ensemble learning · Convolutional neural network · Random forest · Hidden Markov model · Voting classifier

C. Salamea-Palacios (✉)
Universidad Politécnica Salesiana, Cuenca, Ecuador
e-mail: csalamea@ups.edu.ec

J. Guaña-Moya
Pontificia Universidad Católica del Ecuador, Quito, Ecuador
e-mail: eguana953@puce.edu.ec

T. Sanchez · X. Calderón · D. Naranjo
Escuela Politécnica Nacional, Quito, Ecuador
e-mail: tarquino.sanchez@epn.edu.ec

X. Calderón
e-mail: xavier.calderon@epn.edu.ec

1 Introduction

The analysis of cough sounds through artificial intelligence has gained enormous attention in recent years. The study focuses on the automatic detection and classification of cough sounds from among other environmental sounds. One of the problems in this task lies mainly in the insufficient amount of relevant cough sounds compared to different environmental sounds. In many cases, obtaining cough sounds involves ethical issues, the privacy of the participants, the confidentiality and security of the information, and the cleanliness and quality of the data [1–3].

Researchers used to apply data balance methods as oversampling or undersampling or directly through the creation of artificial data through data augmentation procedures to solve these inconveniences [4–6].

Among the ways to analyze the characteristics of cough, the extraction of features has been addressed through the generation of Mel Frequency Cepstral Coefficients (MFCC) and linear predictive coding. This approach allows the recognition of sound data to be transformed into a visual recognition task. In addition, the mentioned characteristics allow representing the data in a spectrogram of the acoustic signals, and the preprocessing can be carried out from a visual approach. These procedures have been widely used in language recognition settings [7–9].

The use of image features in acoustic data detection and classification is another approach in the feature engineering process. For example, the data represented by the image spectrograms of the acoustic signals recorded in the data acquisition can be preprocessed before feature extraction.

Among the models that have been designed to perform cough sound recognition tasks are machine learning and deep learning algorithms. Thus, [10] implemented a classification algorithm based on k-nearest neighbors; some authors implemented support vector machines, [9] modeled multilayer perceptron, [3] used hidden Markov models, and [11] implemented convolutional neural networks.

Convolutional neural network learns input–output relationships based on convolution operations and involves convolution, pooling, and fully connected layers [11, 12]. Random forest is an assembly method that uses a set of decision trees as predictors. The assembly method generally employs bootstrap aggregating, to train different decision trees with different samples with replacement [13].

On the other hand, the hidden Markov models HMM are based on the probabilities of occurrence of well-defined stages within the behavior of processes that generally vary with time or situations and on indirect observations of the same. These stages or states are not necessarily visible or measurable directly, but through their effects or external observations. Transitions between states occur with specific probabilities, which are known, either by learning the probability distribution function that describes such a process or simply by counting the occurrence of such observations and state transitions. Once the HMM is defined, various classification algorithms based on these probabilities can be applied [3, 14].

Nowadays, a trend known as ensemble learning proposes to build a new model from previously trained, similar, or completely different models to obtain a synergistic predictive capacity, that is, better than the predictive capacity of its individual constituents [15–17].

In this work, we propose an ensemble learning approach, through the modeling of a convolutional neural network, a random forest, and a classifier based on hidden Markov models and its subsequent assembly in a hard voting classifier algorithm, for the differentiation of the cough sounds from other environmental events by extracting characters from MFCCs.

2 Methodology

2.1 Database

We have used three datasets in our study. The first one consists of 266 recordings of cough collected telematically and anonymously on the Web site <https://databasecovid19.ups.edu.ec/>. We have extracted, checked, and manually labeled all of the recordings with any cough sound. The second dataset was the [1] database, and we only considered the 54 cough recordings of users who reported dry or wet cough in this case. Finally, the third dataset was the [18] database, consisting of 2000 recordings of environmental sounds, 40 of which were cough sounds.

2.2 System Architecture

Two components were defined for the system architecture: front-end and back-end. In the front-end, the sounds are preprocessed. First, we carried out a data augmentation stage by frequency shift where we varied the pitch of the original audio sample by a random variation factor adjusted to a maximum of 20% [4, 6]. Next, we normalized the amplitude and removed extended silences fragments in each recording. We defined the criteria for carrying this task out heuristically and by visual validation [19]

After this first preprocessing step, the recordings lasted between 0.3 and 4.55 s, with a modal value of 2 s. Consequently, we standardized the duration of each sample to the modal value so that samples lasting more than 2 s were cut into 2-s segments. In contrast, samples lasting less than 2 s were replicated and attached until reaching 2 s. We reanalyzed the resulting samples to discard all the silent ones. This process also modified the ratio cough sound/non-cough sound to 1:4.

On the last stage of the front-end, we carried out a feature extraction step by transforming the samples into Mel spectrograms (MFCC), since the resulting coefficients

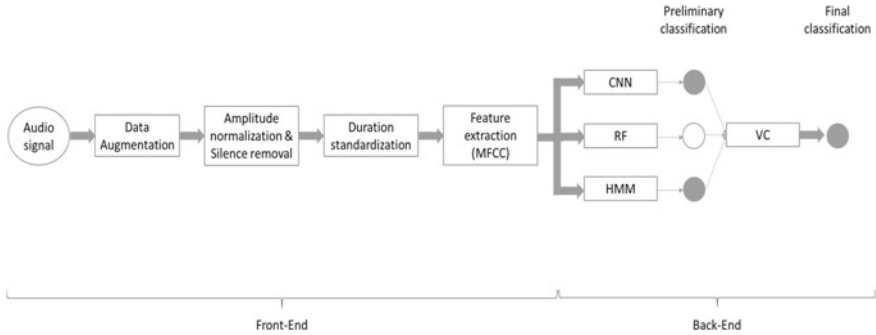


Fig. 1 System architecture

are widely used for speaker identification from environmental sound events as well as speech recognition [20–22].

In the back-end, MFCC were conveniently reshaping to fit the requirements for each classifier modeled in this study. We trained 4 classifiers: a convolutional neural network (CNN), a random forest (RF), a hidden Markov model-based classifier (HMM), and a voting classifier (VC).

We trained the CNN with two convolutional layers, two fully connected layers, and a binary classification layer. Both convolutional layers contain 16 rectified linear units each. The first convolutional layer takes the 16×200 reshaped MFCC as inputs to a 9×3 size filter, followed by a 2×1 layer of max-pooling. The second convolutional layer has a filter size of 3×3 followed by another 2×1 layer of max-pooling. The fully connected layers contain 256 units each. Also, we implemented batch normalization and 0.5-rate dropout regularization to reduce overfitting problems [11, 23]. We developed this model with Keras library.

For the RF classifier, we reshaped the MFCC as a 1-D array and defined 100 trees in the forest, Gini impurity, and the default hyperparameters of the random forest classifier of scikit-learn library [16, 19].

The HMM was built on a class we recreated from hmmlearn of scikit-learn library. We trained a classifier based on hidden Markov model with Gaussian emissions and defined the number of hidden states by maximizing the likelihood [13, 14, 24].

Finally, in Fig. 1, we ensemble the three classifiers (CNN, RF, and HMM) to build a hard voting classifier, which means the VC classifies the data based on the mode of all the three classifiers [17, 25].

2.3 System Performance

We measured the system performance through the accuracy that is the fraction of predictions that the model made correctly; the precision that is the ratio between the cough sounds correctly classified as cough sounds and the instances classified

as cough sounds; the recall that is the ratio between the cough sounds correctly classified as cough sounds and the instances that are actually cough sounds; and F1-score, which is the harmonic mean between precision and recall [26, 27]. In some cases, these metrics give a specific idea about the system behavior. However, in this case, because of the difficulty of the task, it has been necessary to include the confusion matrix as metric to extract information about false-positive and false-negative values, looking for the best configuration that allows to obtain a stable and confidence system.

3 Results

Figure 2 shows that the model tends to stabilize as the iterations progress. We observe that the value of the four performance measures stabilizes at values greater than 0.95 in the validation set.

Table 1 shows that the CNN model in the test set achieved relatively high precision and recall values even when the data were not completely balanced. These results, in turn, translate into a significant value of F1.

On the other hand, Table 2 shows the confusion matrix of the RF model. We observe similar performance to the CNN model: high accuracy, perfect precision,



Fig. 2 CNN accuracy, precision, recall, and F1 score in train and validation sets

Table 1 CNN confusion matrix

		Predicted class		
		Not-cough	Cough	Total
Actual class	Not-cough	208	2	210
	Cough	14	54	68
Total		222	56	278

Table 2 RF confusion matrix

		Predicted class		
		Not-cough	Cough	Total
Actual class	Not-cough	210	0	210
	Cough	15	53	68
Total		225	53	278

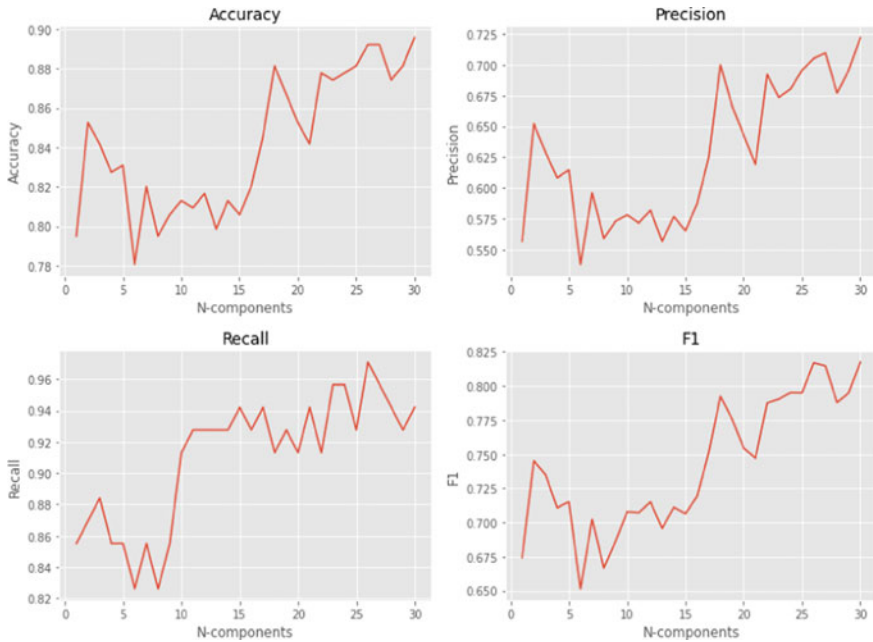


Fig. 3 HMM accuracy, precision, recall, and F1 score in validation set

high recall value, and consequently high F1 score. Additionally, these were very interesting since the RF model was relatively simple, we did not require fine-tuning, and we practically worked with the default values of the hyperparameters.

Figure 3 shows the evolution of the performance metrics as a function of the number of hidden states (n -components) of the HMM model. We can see that there is a growing trend for all metrics. However, before choosing an infinite number of hidden states in the HMM, it is necessary to consider that the greater the number of components, the greater the complexity of the algorithm, and the greater the computational resources.

We train an HMM with 26 hidden states as a first approximation. As shown in Table 3, we obtained a model with a relatively high classification accuracy, yet an unacceptable precision value. However, it is the model with the greater recall compared to CNN and RF.

Table 4 presents the results of the voting classifier ensemble model. This model collects the characteristics of the three previously shown models, and consequently, we obtain a model with balanced values in both false positives and false negatives.

Table 5 shows a comparison of the performance metrics of the models we trained. We observe that the CNN and RF models achieved the highest precision values and the HMM model presents the highest recall value. While the VC model, being assembled with the previous three, achieved the highest performance metrics, so it would be the most suitable model to identify cough sounds.

Table 3 HMM confusion matrix

		Predicted class		
		Not-cough	Cough	Total
Actual class	Non-cough	164	46	210
	Cough	6	62	68
	Total	170	108	278

Table 4 VC confusion matrix

		Predicted class		
		Not-cough	Cough	Total
Actual class	Not-cough	210	0	210
	Cough	14	54	68
	Total	224	54	278

Table 5 Performance of the trained models

Measure	CNN	RF	HMM	VC
Accuracy	0.942	0.946	0.813	0.950
Precision	0.964	1.000	0.574	1.000
Recall	0.794	0.779	0.911	0.794
F1	0.871	0.876	0.705	0.885

4 Discussion

The present work aims to identify the cough sounds of other sound events in the environment. One of the main challenges we faced in this work was the small number of cough recordings available compared to other environmental noises. For this, we used data augmentation in the cough sounds that we had. In addition, the standardization of the duration will obtain a relatively less unbalanced dataset.

The models we trained work on finding patterns underlying MFCCs that differentiate cough sounds from environmental sound events. All the models reached classification accuracies greater than 80%. However, starting from a dataset that is not entirely balanced, we focused our comparison on precision, recall, and, consequently, F1-score.

The CNN and RF models reached the highest precision values. Looking at the respective confusion matrices, we see that, indeed, they are the models with the least number of false positives. These results imply that virtually all sounds predicted as coughing by the CNN and RF models are actually coughing. Thus, if the application objective of this system is to filter the cough data from other sounds and then perform a cough sound analysis, the CNN and RF models are the most suitable for this task [27].

On the other hand, the HMM model is the model that achieved the highest recall. By observing the confusion of the HMM model, we verify that it is the model with the fewest false negatives. This result implies that most of the sounds predicted by the HMM model as not-cough are actually not cough. Thus, if the objective of applying the system is to filter to differentiate cough sounds from other environmental sounds and at the same time lose the least possible number of cough sounds, the HMM model is optimal to perform the classification [27].

The VC model achieved high precision results and recall values similar to CNN. Furthermore, VC is the model that achieved the highest F1-score value. That is, it is the model that had the best balance between precision and recall. Thus, VC could be the optimal model to effectively filter cough sounds from other ambient noises, with a relatively small number of false positives and false negatives. Likewise, the performance metrics shown in Table 5 are similar and even higher than those obtained in other works as the one carried out by [3, 9–11].

Among the reasons we obtained these encouraging results, we found that the data processing and the extraction of characteristics through the MFCC allow us to approximate the task of recognition of cough visually, which has been stated as an optimal approach for this classification task [21, 22]. In addition, we can get rid of non-informative traits such as silent fragments. Also, the ensemble learning approach makes it possible to compensate for the weaknesses of some of the models with the strengths of others, resulting in a much more robust and reliable model [16].

5 Conclusions

In the present work, we assembled a hard voting classifier model with a convolutional neural network, a random forest, and a classifier based on hidden Markov models to differentiate cough sounds from other environmental sound events from the extraction of characters from the MFCC. The results obtained experimentally show that the joint learning approach allows us to solve the weaknesses of some models with the strengths of others. We brought a robust model, with high qualification precision, as well as high precision and recovery.

Among the limitations underlying this work, we find the insufficient number of recordings of cough sounds compared to other environmental sounds. To solve this problem, we apply data augmentation; however, there are additional data balancing techniques that we could use, such as oversampling.

On the other hand, future work could develop new models assembled with more complex algorithms that implement deep learning or transfer learning. In this way, cough detection systems are designed in this way that can be applied as the first stage in a field of growing interest, given the current situation arising from the COVID-19 pandemic.

Acknowledgements To the Corporación Ecuatoriana para el Desarrollo de la Investigación y Academia, CEDIA, for the financing provided to research, development, and innovation through the CEPRA projects, especially the project CEPRA-XV-2021-011: Caracterización de la tos provocada por el COVID-19 en pacientes con diagnóstico positivo.

The authors thank Escuela Politécnica Nacional, Universidad Politécnica Salesiana, and Pontificia Universidad Católica del Ecuador.

References

1. Brown, C.H.: Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. In: KDD'20, pp. 3474–3484 (2020). Obtenido de. <https://doi.org/10.1145/3394>
2. Gibson, P.G.: Speech pathology for chronic cough: a new approach. *Pulmonary Pharmacol. Therapeut.* **22**(2), 159–162 (2009). Obtenido de. <https://doi.org/10.1016/j.pupt.2008.11.005>
3. Matos, S.B.: Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Trans. Biomed. Eng.* **53**(6), 1078–1083 (2006). Obtenido de. <https://doi.org/10.1109/TBME.2006.873548>
4. Cui, X.M.: Data augmentation for deep neural network acoustic modeling. *IEEE/ACM Trans. Audio Speech Lang. Process.* **23**(9), 1469–1477 (2015). <https://doi.org/10.1109/TASLP.20>
5. García, M.A.: Data augmentation para la Clasificación Automática de la Calidad vocal data augmentation in automatic classification of voice quality. In AJEA (2020) <https://doi.org/10.33414/ajea.5.748.2020>
6. Shorten, C.: A survey on image data augmentation for deep learning. *J. Big Data* **6**(60), 1–48 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
7. Arturo, G.M.: Reconocimiento de voz basado en MFCC, SBC y Espectrogramas. *Ingenius* 12–20. (2013)
8. Grama, L.: Choosing an accurate number of mel frequency cepstral coefficients for audio classification purpose. In: International Symposium on Image and Signal Processing and Analysis, ISPA, Ispa, pp. 225–230 (2017). <https://doi.org/10.1109/ISPA.2017>

9. Mitra, V.: Content based audio classification: a neural network approach. *Soft Comput.* **12**, 639–646 (2008). <https://doi.org/10.1007/s00500-007-0241-4>
10. Khunarsal, P.L.: Very short time environmental sound classification based on spectrogram pattern matching. *Inf. Sci.* **243**, 57–74 (2013). <https://doi.org/10.1016/j.ins.2013.04.014>
11. Amoh, J.: Deep neural networks for identifying cough sounds. *IEEE Trans. Biomed. Circ. Syst.* **10**(5), 1003–1011 (2016). <https://doi.org/10.1109/TBCAS.2016.2598794>
12. Bouwmans, T.J.: Deep neural network concepts for background subtraction: a systematic review and comparative evaluation. *Neural Netw.* **117**, 8–66 (2019). <https://doi.org/10.1016/j.neunet.2019.04.024>
13. Hoffait, A.S.: Early detection of university students with potential difficulties. *Decis. Support Syst.* **101**, 1–11 (2017). <https://doi.org/10.1016/j.dss.2017.05.003>
14. Teyhouee, A.: Cough detection using hidden markov models. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11549 LNCS, pp. 266–276 (2019)
15. Deng, H.R.: A time series forest for classification and feature extraction. *Inf. Sci.* **239**(February), 142–153 (2013). <https://doi.org/10.1016/j.ins.2013.02.030>
16. Mohammed, E.A.-N.: An ensemble learning approach to digital corona virus preliminary screening from cough sounds. *Sci. Rep.* **11**(1), 1–11 (2021). <https://doi.org/10.1038/s41598-021-95042-2>
17. Ruta, D.: Classifier selection for majority voting. *Inf. Fusion*, **6**(1), 63–81 (2005). <https://doi.org/10.1016/j.inffus.2004.04.008>
18. Piczak, K.J.: ESC: Dataset for environmental sound classification. In: *MM 2015—Proceedings of the 2015 ACM Multimedia Conference*, pp. 1015–1018 (2015). <https://doi.org/10.1145/2733373.2806390>.
19. Pahar, M.K.: COVID-19 cough classification using machine learning and global smartphone recordings. *Comput. Biol. Med.* **135**, 1–13 (2021). <https://doi.org/10.1016/j.compbiomed.2021.104572>
20. García, M.A.: Power cepstrum calculation with convolutional neural networks. *J. Comput. Sci. Technol.* **19**(2), 132–142 (2019). <https://doi.org/10.24215/16666038.19.e13>
21. Grama, L., Rusu, C.: Choosing an accurate number of mel frequency cepstral coefficients for audio classification purpose. In: *International Symposium on Image and Signal Processing and Analysis, ISPA, Ispa*, pp. 225–230 (2017). <https://doi.org/10.1109/ISPA.2017>
22. Martínez Mascorro, G.A., Aguilar Torres, G.: Reconocimiento de voz basado en MFCC, SBC y Espectrogramas. *Ingenius* (2013). <https://doi.org/10.17163/ings.n10.2013.02>
23. Lecun, Y.B.: Gradient-based learning applied to document recognition. *IEEE Proc.* **86**(11), 2278–2324 (1998). <https://doi.org/10.1016/j.bbrc.2005.03.111>
24. Jeebun, S.: Optimal number of states in hidden Markov models and its application to the detection of human movement. *Univ. Mauritius Res. J.* **21**, 438–469 (2015)
25. Żabiński, G.G.-J.: Multi-classifier majority voting analyses in provenance studies on iron artefacts. *J. Archaeol. Sci.* **113** (2020). <https://doi.org/10.1016/j.jas.2019.1>
26. Chicco, D.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **21**(6), 1–13 (2020). <https://doi.org/10.1186/s12864-019-6413-7>
27. Juba, B.: Precision-recall versus accuracy and the role of large data sets. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 4039–4048 (2019). <https://doi.org/10.1609/aaai.v33i01.33014039>

Advantages and Benefits of Big Data in Business Communication



Mario Román Rivera and Karina Valarezo González

Abstract Big data integrates and evaluates large amounts of data collected by a company, which when crossed, allow obtaining indicators that contribute to visualize and improve organizational management in the internal and external environment of a business. The area of communication, marketing and advertising knowledge has been widely benefited by the use and exploitation of data, makes it possible to make intelligent decisions when proposing strategies that are better targeted and adapted to a specific target audience and to objectives specific to the organization, among others, to attract and/or retain consumers or customers; and, evaluate trends and new business opportunities—innovation—to ensure the sustainability of the business. Working under a data mining approach brings additional value and is that the actions implemented can be evaluated in real time or very quickly to improve their effectiveness. The main objective of the study is to demonstrate the criteria according to the experience of those responsible for communication and marketing of companies in Ecuador according to the ranking of the magazine Ekos Business 2019 and the opinion of experts in big data and communication, on the advantages and benefits of big data in the management of the strategic communication of organizations. The results show that, in Ecuador, the use of big data is being adopted for the development of better business and strategic communication strategies.

Keywords Big Data · Communication · Business

M. R. Rivera (✉)
Universidad Técnica Particular de Loja, Loja, Ecuador
e-mail: meroman3@utpl.edu.ec

K. V. González
San Cayetano Alto, C. París, Loja 110107, Ecuador

1 Contextualization

1.1 *Big Data, Definitions and Applicability in Strategic Communication 360*

Since the massive use of the Internet, organizations have been forced to transform the way information is stored and analyzed, “organizations have reached a point where they don’t know how much data they will generate and how to process it properly” [1].

Big data, as its name implies, the management of a large amount of data, macro data, structured or not, so large/voluminous and complex that are difficult or impossible to process with traditional methods, is information that we may have or collect as individuals or organizations.

The term Big Data not only refers to the volume of data, but to the technologies associated with the capture, administration and visibility of these; and not only the volume but the variety of these and the speed of access and processing are taken into account. [2, p. 312].

This data mining cannot be processed through common programs with basic computer skills, but requires professionals or specialized human talent, special software and machines with a large memory and processor capacity.

The large amounts of information currently generated are presented in “different levels of complexity, which cannot be processed by commercial tools or common technologies that can be applied to a traditional database” [3].

The attributes of big data according to Martínez-Martínez and Lara-Navarra [4]. They are (volume, speed, variety, value) established as an important axis of attention for the study, processing and processing of the data collected through emails, videos, documents, text messages, labels, social media among others.

Meanwhile, Revuelta [2] adds three more V’s to the concept of data:

- Feasibility: entities must be realistic in the technical and human needs necessary to provide effective use of data, establishing customized resources in each project.
- Data visualization: it is essential to use visualization tools to be able to reach an understanding of the data.
- Data value: Data should be transformed into information that is useful for an entity’s decision-making. The processing of this data must therefore be properly directed [2, p. 313]

Big data has become the basis for companies, among other things, it offers companies to know the behavior, concerns, the profile of their audience, their geolocation, etc., information that is often unknown in the company; that is, it is a reference of what the company should know and do with the large amount of data that it generates itself. The data can be analyzed in such a way that the company identifies not only the problems, but also the possible solutions for these and of course their strengths to take even greater advantage of them. Every day, companies are faced with public valuation, demanding better results in terms of product/service quality, care of the

environment and contribution to the environment in which they operate; therefore, having information can make the difference between one organization and another.

The receiver today not only listens to what the brand tells him, he has become a propositional and demanding actor at the same time, “To be a recognized and preferred brand by the consumer it is necessary to differentiate and be in the top of mind and also be the first option considering that the brain has the capacity to store only 7 brands” [5, p. 17]. A number of techniques and methods have now emerged that have changed the management of strategic communication in all areas in order to offer new alternatives that convert the proposals of the most interesting brands and organizations (Table 1).

In short, we can analyze how communication has evolved, combining even with other branches of technical and analytical nature, In addition, it leaves the door open for us to know what we can expect for the future of one of the most multi-functional branches of the social sciences.

1.2 Impact of Big Data on Communication 360

Data analysis and management is not something new in the management of 360 communication where they are included: public relations, advertising and marketing in organizations. The traditional models of strategic communication are formed by different phases of analysis and constant evaluation that have implicit the analysis of data. However, due to the steady growth of new technologies, the need to implement more flexible and cross-cutting planning and control processes in the research phases to overcome traditional approaches is raised [7].

As companies increase their presence on the web and, especially, on social networks, it is essential to know the possibilities offered by social analytics, including calculating the successes and failures of organizations, detecting possible criticism, define positioning, etc. [8, p. 125]

Just as communication evolves because of its new trends, business marketing also converges in dynamic growth, it is vital to understand that “The evolution of marketing and communication is encompassed in three stages: marketing and communication 1.0, marketing and communication 2.0 and marketing and communication 3.0. The latter is also known as Social Marketing” [8, p. 124].

Thanks to the Internet, we are currently in the third Industrial Revolution, and in the first phase of the fourth, because it is based on the iconic invention of this (Internet) “and is characterized by a fusion of different technologies (artificial intelligence, robotics, machine learning...) that blurs the line that hitherto separated the spheres of the physical, the digital and the biological” [8, p. 117].

Human beings create more and more information every day, all this information comes from different places: social networks, Smartphones, Smart Cities, companies, so-called M2M communications (machine to machine), digital sensors whether electrical or temperature measurement, seismic, it is estimated that there are more than 30 million interconnected sensors in different sectors and this number is expected to grow annually by 30% [9: 13].

Table 1 New trends in communication

Communication 360°	Based on the combined use of all possible tools and methods to communicate with the target audience
Online advertising and Web presence	Stands out for the formats for advertising
Internet as a means of communication	Intervening channels, platforms or digital media
Viral communication and buzz marketing	They rely on “mouth-ear” communication by taking advantage of rumors that spread quickly and exponentially
Communication techniques on mobile devices	It refers to issues related to mobile networks: Webs, mobile advertising formats and other messaging formats (SMS-WhatsApp) subscription services and advertising formats
New trends in internal communication	Refers to communication processes and mechanisms within organizations
Content communication and marketing	Refers to the characteristics and elements that make up content marketing
Advergaming as an advertising tool	It consists of including a brand or product in a very particular medium: a video game
Communication integrated in the environment: street marketing and ambient marketing	They stand out because they open the doors to new prospects for advertisers and consumers
Management of experiential and sensory communication	Focuses on creating experiences and using the senses
Neuromarketing and communication	One of the currents that has raised the most expectations in understanding the internal mechanisms of the consumer
Tryvertising and advertorials	It is based on advertising by testing the product and combining advertising and publishing
Big data and communication	It is based on the analysis of the large amount of data available today especially generated by the activity of individuals on the Internet for decision-making

Note Adapted from Pintado and Sánchez [6]. New trends in strategic communication

In this sense, the use and application of big data open the doors to a great opportunity in strategic communication for the potential and opportunities it offers when making decisions and anticipating new possible changes in the market or public needs “to create better objectives, more strategic positioning and target audiences, and to discover more opportunities to quantify and improve the contribution of public relations to business success” [10, p. 15].

Big data allows the communication and marketing teams of organizations, the real-time analysis of large amounts of information and thanks to them to anticipate

the creation of new services and/products; but it also poses challenges and opportunities; for example, how to address data, its importance in decision-making and how it affects business transformation, new skills and roles, data monetization, organizational changes and new business models. These predictions are useful for making strategic decisions in the organizational structure.

Today, companies present a high demand for technologies to solve many of their problems or business situations. The transversality and interaction between communication and marketing departments is becoming a necessity in order to stay at a competitive level in a rapidly changing environment [11, 12]. Thus, big data is beginning to be considered as an option to obtain insights about the public or the context in which the company develops.

As a relatively new trend, it is important to know the current state of use of big data in Ecuador to be able to create an academic background that allows generating a contribution to the line of research related to the use of big data in Ecuadorian companies and its incidence in departments as in the communication and its strategic management.

2 Methodology

Within a research work, it is important to manage criteria such as objectivity and systematicity that, according to Maya [13], are bases of a scientific thought that contribute to the development of a research work “to understand the research process and the results thereof”. Maya [13, p. 14], taking into account the analysis of studies carried out without neglecting new observations, discoveries and methodologies are applied.

For the development of this work, a mixed methodology was used; the quantitative one, based on probabilistic induction and representation of data through numbers and percentages, is objective, confirmatory, oriented to obtaining solid results through reproducible processes [14]; and the qualitative one, which can be defined as “research that produces descriptive data: people’s own words, spoken or written, and observable behavior” [15: p. 7].

As information collection tools, an online survey was used which consists of a structured questionnaire or set of closed questions or multiple choice, allowing to obtain information on a specific topic, in a population from a sample [16]. For the application of this survey, a convenience sampling was taken as a reference for 44 companies in the business ranking of the magazine *Ekos* 2019 [17], of which 14 agreed to participate by completing the questionnaire of 35 questions aimed at learning about three important aspects: big data in the management of strategic communication, importance and benefits of big data in communication.

In addition, a structured interview was applied, which allows obtaining perceptions from experience [18]; for the present study, a questionnaire of 22 questions was applied to experts in strategic communication and big data applied to communication, whose names are described in Table 3, in order to collect data complementary

to the online survey and thus to carry out a systematic comparison and analysis. This interview was intended to know from the academic point of view and the experience, the competitive advantages of big data in communication management.

The techniques implemented, the information collected and the analysis of it, made it possible to achieve the objective that the study set out: to analyze the criteria of those responsible for communication and marketing, of the best ranking companies in Ecuador and of experts in strategic communication on the development, application, advantages and benefits of big data, objective that was dismissed from the following question: Does the application of big data in the communication management of communication strategies and brand positioning of companies with better ranking in Ecuador generate advantages and benefits to the organization?.

3 Results

In order to improve the organization and understanding of the results obtained, they have been divided into two sections (Table 2):

3.1 *Business Environment and Professionalization*

Business Area of the Company and Professionalization of Respondents

The sector with the greatest participation of professionals who completed the survey is linked to the banking sector (21.43%), followed by the marketing of products either of primary need or related, which is strongly related to that of food (both have 14.29%) (Figs. 1 and 2).

With regard to specialization, it is important to note that the majority of respondents have studies related to strategic (28.57%) and organizational communication (21.43%), with a training in communication sciences, digital communication or marketing (14.29%) and finally professionals with studies in big data analytics (7.14).

It is also important to note that in the area of professionalization, 50% of those surveyed completed their third level studies in the field of communication sciences, 35.71% in marketing or journalism (14.29%). In addition, the majority of participants

Table 2 Classification of results

Business environment and professionalization	Use and importance of big data in communication management
The area of business in which communication professionals work is completed and the degree of specialization according to their professional training	The contribution, advantages and benefits of big data in the communicational structure of the organization for decision-making are analyzed

Table 3 Advantages of big data in communication management, according to the criteria of communication experts

Experts	Concepts
Ph.D. Carlos Toural	It allows to know more and better each and every one of the contexts, agents and processes that operate in the strategic communication
Ph.D. José Miguel Túñez	It allows everything. It allows the public to be loyal
Ph.D. María Dolores Brito	The digital use of users can be accurately and accurately measured
Ph.D. Gustavo Cusot	It allows us to understand how I make decisions and how I adjust the strategic plan that I am implementing
Ph.D. Martha Lucía Buenaventura	Big data allows to have governance. It generates fewer errors, maintaining the privacy and security of the organization
Ph.D. Susana Cadena	It no longer gives us a competitive theme; it gives us a survival theme
Ph.D. Albertina Navas	Makes it possible to stop being operational and move to an analytical level to make decisions based on those data
Ph.D. Marcelo Santos	Speed. Accuracy
Ph.D. Carlos Arcila	Real-time access to public perception or audiences that help improve organizations' communication strategies
Ph.D. Alejandro Álvarez	Enables organizations to be more resilient to constant change

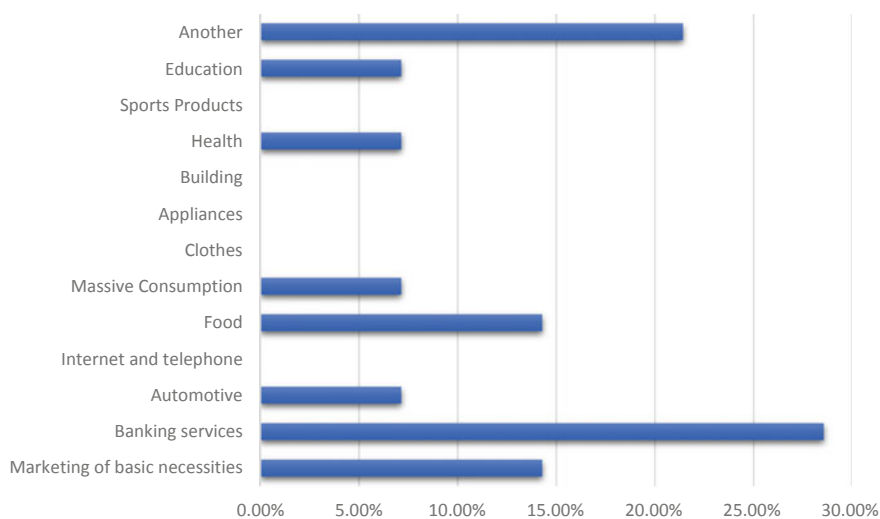


Fig. 1 Business area of the company

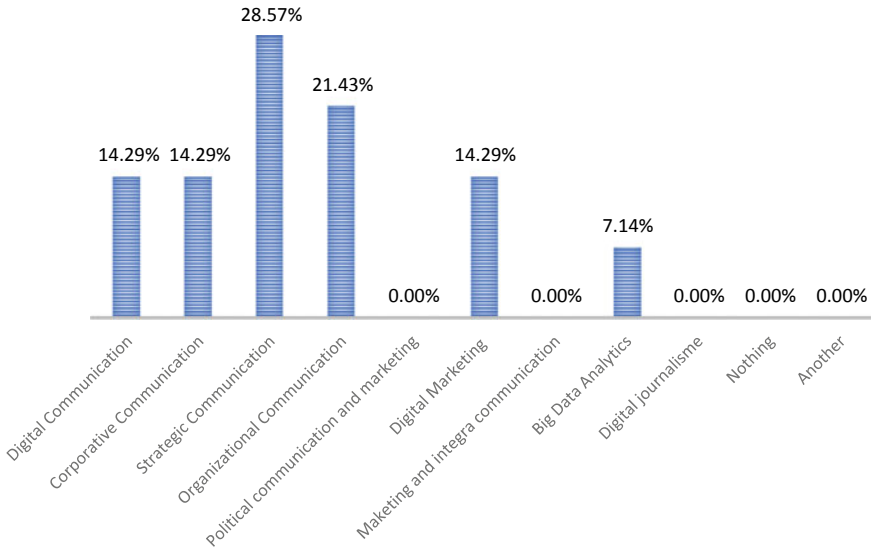


Fig. 2 Communication officer’s area of professional specialization

(71.43%) have a fourth level specialization and the remaining percentage (28.57%) only third level studies.

On the other hand, it is also noted that according to the database collected for the implementation of the online survey it was evident that most of the people responsible for communication management in companies are male, Data that, at the time of the implementation of the survey was ratified with a participation of 57.14%, compared to 42.86% of female professionals.

3.2 Use and Importance of Big Data in Communication Management

Contribution, Benefits and Advantages

According to communication and marketing professionals consider that Big Data has an incidence (75%) when creating and executing communication strategies. It is important to have an acceptable understanding of this type of strategy, especially if the organisation in which it will be applied on a daily basis tends to generate large amounts of information such as banks or other companies, therefore, according to the study, 50% of those responsible for communication are very familiar with the use of Big Data or are completely familiar with its application (14.29%) and also use it in their institution, However, it is still striking the fact that there are still companies (35.71%) that their knowledge about the use of this trend is low, since they are somewhat familiar and therefore do not apply it in their organization (Figs. 3 and 4).

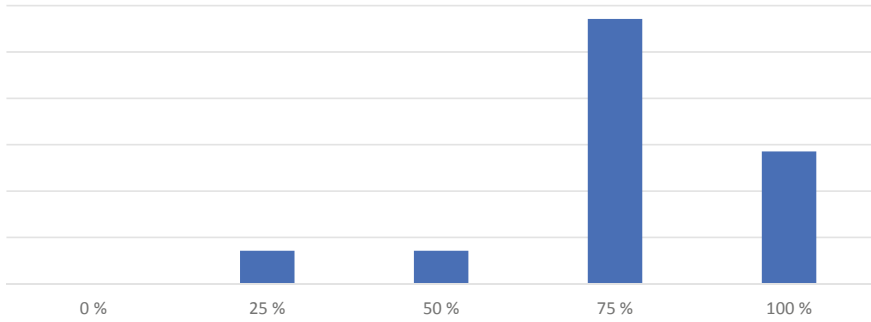


Fig. 3 Contribution of Big Data to the company’s communication and marketing strategy

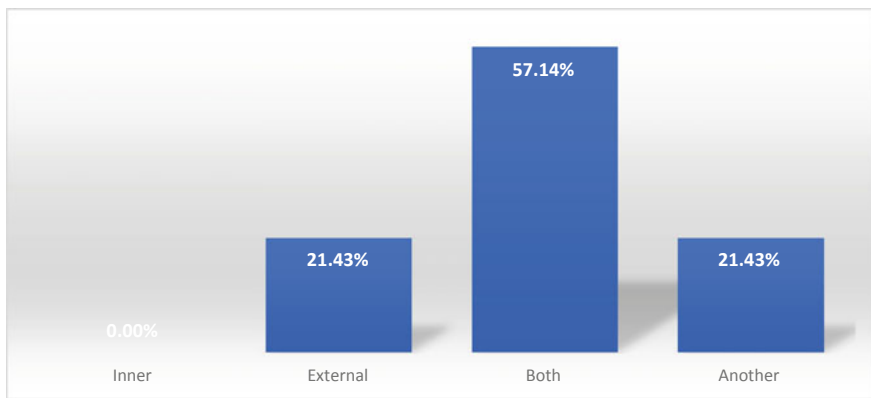


Fig. 4 Big Data in communication strategies

Regarding the applicability of big data in communication management, professionals (57.14%) point out that it is applied both in internal and external communication actions, unlike another sector (21.43%) who points out that it applies only to external management and others state that in none.

From the responses of the experts interviewed, the most significant advantages, benefits and contributions that can be obtained with the application of big data in communication were collected, as opposed to conventional strategies (Fig. 5; Table 3).

Regarding the most important benefits of big data in the communication management of companies in Ecuador, it is noted that the main ones are evaluate trends, geolocate consumers and improve sales (71%), added to new business opportunities (57%) and attracting new consumers and followers on social networks (50%).

Data-based communication is a competitive advantage, according to Cadena (2021), “it is no longer a matter of being better than the other, it is a matter of whether I want to survive with my company or not, that easy. [...] the pandemic has not given us time to react, for example, many companies have gone bankrupt

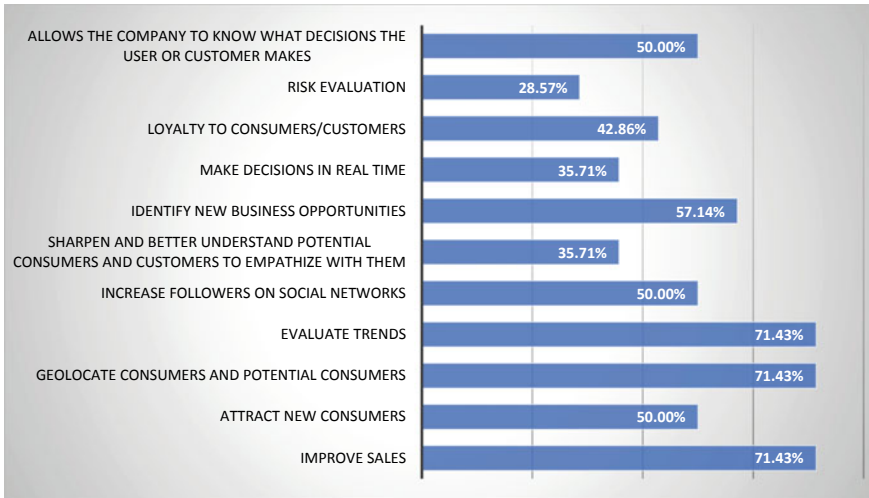


Fig. 5 Benefits of big data in communication management

because they do not risk the digital issue and communication is one of the crucial elements”. Analyzing the responses of Túñez, Taural and Cadena, it is highlighted that the analysis of the data allows to know something in a deeper way and allows to make better decisions within the strategic communication. A strategy can be creative, but if it is based on weak research, it can risk its effectiveness.

3.3 Indicators and/or Variables to Be Taken into Account for the Construction of the Communication Strategy

The indicators are born based on the objectives set; Below are those that, in the opinion of (Cadena, Taural, Brito, Cusot, Arcila and Santos, online interview, March 2021) should be considered from the design of the strategy to achieve the expected impact.

(Fig. 6).

The survey also consulted on the receptivity of users to the content produced based on the contributions provided by big data. Forty-two percent mentioned that their followers showed very good acceptance and 35% indicated that the receptivity was excellent and 21% stated that it was good, of which it can be established that big data gives 98 percent reliability over conventional strategies.

¿When to Apply Big Data?

Finally we close by analyzing when it is pertinent to apply big data and when it is not; we will begin by analyzing the when not by the hand of Marcelo Santos, who watches over the integrity of users since “to the extent that one identifies that one can

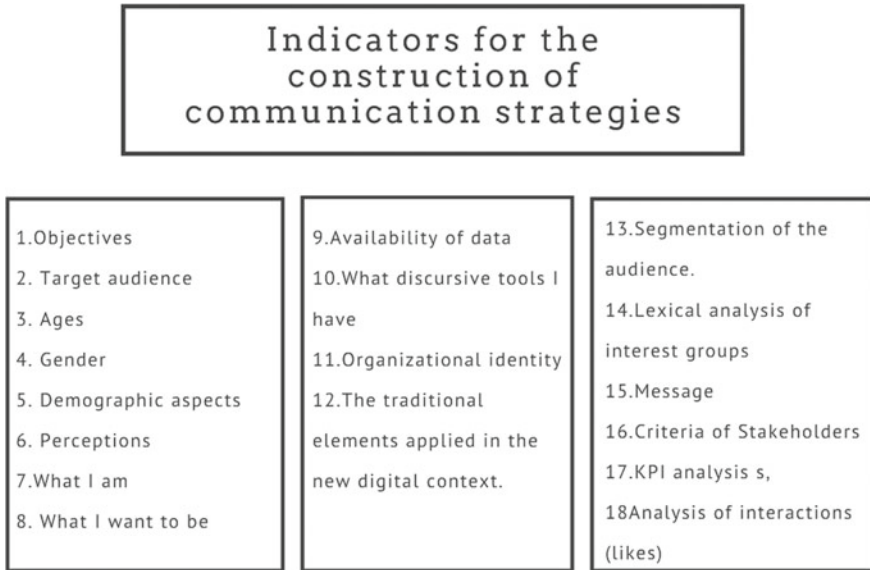


Fig. 6 Indicators for building communication strategies based on big data. *Note* Enumeration of indicators does not define level of importance

harm some person or group of people with the information and that, when exposing certain data I will end up affecting it, in that scenario should not apply” (Santos 2021) complementing what expressed by Túnuez (2021) should not be applied, “when it does not contribute, when they are very small elements”.

And when yes, “when there are organizations that have a great impact on the lives of many people, And when I do not have time or access to other types of data” everything will depend according to Cadena (2021) “depending on the impact and the results you want to achieve” and also “depending on the type of data, variety and volume, if you have variety and volume is vital”.

For Arcila (2021), Big Data must be applied when the information I require from the data cannot be obtained by a manual process. “If a human can manually access the information you need about the data, do not use big data. What human beings cannot do manually has to be done by big data”. This depends on the volume and complexity of the data to be analyzed.

For its part Buenaventura (2021) according to its experience should always be applied (big data), “data has been key, data is the basis to develop communication strategies”. And that has to say of Brito (2021), much will depend on the steps that are indispensable, according to the campaign that wants to publish”. This is not mathematical, it has certain indicators, but these quantitative indicators need to be complemented with the qualitative.

4 Discussion and Conclusions

It is interesting to know that the largest number of communication directors of the companies surveyed (21.43%) belongs to the banking and finance sector; however, it would come as no surprise that their communication managers are linked to the world of data analysis, as their organizations record and analyses a large amount of data on a daily basis. This action is called big transaction data which “includes data from mass transactions of call centers, banking, finance, customer service, etc.,” [19: 436]. That to the post will be used for the creation in strategies that allow to strengthen the loyalty of its clients internally and externally, generating a competitive advantage in the market before other business sectors.

Big data and communication are related; the impact they have on the creation of strategies and action plans according to companies is 75%. This technology is frequently used to maintain or create brand value, launch products effectively and optimize audience and channel segmentation to create deeper customer relationships [20]. To achieve these actions to give the expected result, it is necessary to have an acceptable knowledge in this field; therefore, it is not by chance that the companies with better ranking in Ecuador, have professionals graduated in the field of strategic communication, digital, corporate, organizational, business, marketing and big data analytics because their knowledge and skills on this type of trends allow them to propose better strategies and generate greater confidence in the organization when interpreting, manage and make visible the data obtained.

Regarding the advantages that big data can provide the communication strategy, two aspects should be borne in mind that the communication experts surveyed point out; the first, that the mere fact of applying data in communication already generates a great advantage over the rest; and, secondly, it is no longer about being better than the rest, in many cases, it is a matter of survival, at the moment “only the big companies are the ones who see the advantages, and some already apply with remarkable success” [21, p. 8]. It is vital that organizations commit to the implementation of these concepts in order to maximize their potential both for faster decision-making and for the implementation of new high-impact strategies.

Organizations in Ecuador are using big data to evaluate trends, improve sales, geolocate consumers, among their main options, which allows them to identify new opportunities to generate business. These indicators have begun to generate positive changes in the management of communication and/or business marketing, since the content for the Web produced based on the contributions of big data has a 42% acceptance and loyalty of users, providing 98% reliability over conventional strategies. Inbound marketing, for example, focuses on user loyalty through the sending of personalized messages, not only according to its ranking as a stakeholder, but according to the level of relationship it has with the company, with the ultimate purpose of loyalty and achieving a purchase [22].

It must be clear that big data is not a magic solution to the communication of the company; it is a tool that should complement and improve the actions, results and objectives of the organization. It is necessary to evaluate when it is useful to apply

it. Within the communication, big data contributes to the development of strategies. It is important to highlight this because the data does not redefine a topic by itself, its interpretation and its comparison with other aspects (environment, internal and external aspects, etc.), are what determine its success.

Data analysis is beneficial to the company, they are part of what is called business intelligence, and that is, “make it possible to take advantage of the storage of large volumes of data provided by management information systems to transform them into value for the company by generating knowledge and making timely decisions in real time” [23, p. 40]. In addition, according to experts bring with them economic returns, competitive advantages, real-time decision-making, knowing and loyalty potential customers and mainly allows to optimize the budget.

Finally, it is important to note that for the development of this research and application of data collection instruments. There were several limitations related to the measures of social isolation and telework caused by the COVID pandemic-19, which made it difficult to approach the public objects of study.

References

1. EMC Education Services: Data Science and Big Data Analytics. Discovering, Analyzing, Visualizing and Presenting Data. Wiley, Estados Unidos Indiana (2015)
2. Revuelta Bayod, M.: “Big data”: crisis y nuevos planteamientos en los flujos de comunicación de la cuarta revolución industrial. *Area abierta* **18**(2), 309–324 (2018). <https://doi.org/10.5209/ARAB.59521>
3. Merizalde, M., Ballagán, E.: Aplicación del Big Data como estrategia de negocio en proyectos de emprendimiento en Ecuador, vol 18 No. 18, pp. 114–123. *Revista científica Ciencia y Tecnológica*, Guayaquil, Ecuador. <https://docplayer.es/209861048-Revista-ciencia-tecnologia-no-18-30-de-abril-de-2018-issn-impreso.html> (2018)
4. Martínez-Martínez, S., Lara-Navarra, P.: El big data transforma la interpretación de los medios sociales. *El Profesional de la Información* **23**(6), 575–581. <http://profesionaldelainformacion.com/contenidos/2014/nov/03.pdf> (2015)
5. Rovira, J.: Persona no consumidor. ESIC-Alfaomega, Bogotá, Colombia (2017)
6. Pintado, T., y Sánchez, J.: Nuevas tendencias en comunicación estratégica. ESIC. Editorial, Madrid, España. Recuperado de. <https://cutt.ly/enCnQbj> (2017)
7. Wiencierz, Christian y Röttger, U.: The use of big data in corporate communication. *Corp. Commun: Int. J.* **22**(3), 258–272 (2017). <https://doi.org/10.1108/CCIJ-02-2016-0015>
8. Bustamante, N., & Guillén, S.: Un acercamiento al Big Data y su utilización en comunicación. *Mediaciones Sociales* **16**, 115–134 (2017). <https://doi.org/10.5209/MESO.58112>
9. López, D.: Análisis de las posibilidades de uso de Big Data en las organizaciones. Universidad de Cantabria. <https://repositorio.unican.es/xmlui/bitstream/handle/10902/4528/TFM-DavidLópezGarcíaS.pdf?sequence=1> (2013)
10. Weiner, Mark y Kochhar, S.: Irreversible: the public relations Big Data revolution (IPR White Paper). Institute for Public Relations, Gainesville, FL. Recuperado de. https://instituteforpr.org/wp-content/uploads/IPR_PR-Big-Data-Revolution_3-29.pdf (2016)
11. González, et al.: La estrategia del Big Data como factor clave de la competitividad en las empresas. *Revista Colombiana de Tecnologías de Avanzada* **1**(31), 57–65 (2018). <https://doi.org/10.24054/16927257.v31.n31.2018.2765>
12. Rodríguez, C.: Tendencias en business intelligence del Big data al social intelligence. 57–62 (2017)

13. Maya, E.: Métodos y técnicas de investigación, Distrito Federal, México. D.R. © 2014 Universidad Nacional Autónoma de México. No págs. 1–90. Recuperado de. http://www.librosoa.unam/bitstream/handle/123456789/2418/metodos_y_tecnicas.pdf?sequence=3&isAllowed=y (2014)
14. Del Canto, E., Silva, A.: Metodología Cuantitativa: Abordaje Desde La Complementariedad En Ciencias Sociales. *Revista de Ciencias Sociales* **0**(141) (2013). <https://doi.org/10.15517/rcs.v0i141.12479>
15. Quecedo, L., Castaño, G.: Introducción a la metodología de investigación cualitativa. *Revista de Psicodidáctica* **14**, 5–40 (2002)
16. Hueso, A., Cascant, M.: Metodología y Técnicas Cuantitativas de Investigación. Universitat Politècnica de València. https://riunet.upv.es/bitstream/handle/10251/17004/Metodologíafayténcicascuantitativasdeinvestigación_6060.pdf (2012).
17. Ekos Negocios: Ranking Empresarial del Ecuador. Ekos. <https://www.ekosnegocios.com/ranking-empresarial> (2019)
18. Vargas, I.: La entrevista en la investigación cualitativa. *Calidad en la Educación Superior* **3**(1), 119–139 (2012)
19. Monleon-Getino, A.: El impacto del Big-data en la Sociedad de la Información. Significado y utilidad. *Historia y Comunicación Social* **20**(2), 427–445. https://doi.org/10.5209/rev_hics.2015.v20.n2.51392 (2016)
20. Miquel, S., Aced, C.: Big data: la revolución de los datos y su impacto en la comunicación corporativa. *Comunicación y Hombre* **16**, 115–132 (2020). <https://doi.org/10.32466/eufv-cyh.2020.16.602.115-132>
21. Fernández C.: Cómo las empresas pueden impulsar su negocio a través de las plataformas e-commerce con el big data, el aprendizaje automático y el management científico. Quality Excellence S.L, Madrid, España. <https://url2.cl/RRPbH> (2017)
22. Aguado-Guadalupe, G., García-García, A.: Nuevas métricas de audiencia al servicio del Inbound Marketing. *Doxa Comunicación. Revista interdisciplinar de estudios de comunicación y ciencias sociales*, 81–98 (2018). <https://doi.org/10.31921/doxacom.n26a4>
23. García E, García M (2017) Análisis de los Modelos de Inteligencia de Negocios basados en Big Data en las Pymes del Ecuador, Quito, Ecuador. <https://url2.cl/ghCbx> (2017)
24. Cortés, S., Pedrol, E.: El uso del Big Data en las estrategias de comunicación (2017). *Corporativa*, pp. 1–62. *Revista Uno, Llorente y Cuenca*. <https://outlook.live.com/owa/?path=/attachmmentlightbox> (2016)
25. Merizalde, M., y Ballagán, E.: Aplicación del Big Data como estrategia de negocio en proyectos de emprendimiento en Ecuador vol 18, No. 18 pp. 114–123. *Revista científica Ciencia y Tecnológica*, Guayaquil, Ecuador. <https://docplayer.es/209861048-Revista-ciencia-tecnologia-no-18-30-de-abril-de-2018-issn-impreso.html> (2018)

Estimating Customers' Profitability: Influence of RFM Attributes, Web Metrics and Product Data



Sunčica Rogić , Ljiljana Kaščelan , and Vladimir Đurišić

Abstract This paper presents a method for the evaluation of attribute significance for the prediction of customer profitability. The proposed approach is based on the support vector regression (SVR), compared to most previous studies, which were based on a subjective managerial assessment. The analysis was conducted on two datasets—one with a smaller and another with longer prediction horizon. Results of this study point to Recency Frequency Monetary (RFM) variables as the particular attribute group significant for the profitability prediction, in comparison to the customer response modeling, where customers' web log data are proven to be dominant factors in previous research. Additionally, out of the three purchasing behavior variables, Monetary attribute showed superior influence in short-term profitability prediction, while its significance is almost equated with Recency and Frequency variables in the longer prediction horizon.

Keywords Customer profitability · RFM · Support vector regression · Web metrics

1 Introduction

Direct marketing has gained in importance along with the rapid decline in communication costs. As a result of the low cost of Internet access, the opportunity has opened up for companies to create a direct connection with thousands of consumers in a way that was unthinkable with the use of intermediaries [1]. The further decline in communication costs has also brought certain challenges for direct marketing, taking into account the increasingly competitive environment and lack of customer loyalty, as well as the competition for consumer attention by a growing number of companies, which greatly impacts customer profitability.

Research highlights the importance of being able to identify specific customer segments within general segments and define different marketing strategies based on their patterns of behavior [2]. By gaining the knowledge of customer profitability,

S. Rogić (✉) · L. Kaščelan · V. Đurišić
Faculty of Economics, University of Montenegro, 81000 Podgorica, Montenegro
e-mail: suncica@ucg.ac.me

appropriate strategies can be created and marketing decision-making systems can be improved. Customer profitability analysis provides important information and metrics that are crucial for the allocation of marketing resources, both to individual consumers and to defined segments. Given that modern consumers are extremely demanding and tech savvy, companies have a need to create customized offers, as well as personalized “one-on-one” strategies.

In this paper, customer profitability will be assessed based on Recency Frequency Monetary (RFM) method. Results of RFM analysis can help group customers into segments based on profitability, which can then be further analyzed for future marketing activities. The aim is to present an approach of predicting customer profitability, based on Support Vector Regression (SVR), as well as testing the significance of RFM attributes for this prediction, in comparison to web metrics, customer data and product description data. The significance of RFM attributes is assessed in terms of their influence on the SVR performance, compared to previous research, where the importance for RFM attributes was based on a subjective judgment. The paper also suggests a procedure for assessing the significance of these attributes on a specific customer basis because it can vary significantly depending on the industry.

This paper is organized as follows: the second section provides a summary of relevant literature for the customer profitability prediction based on RFM method. Third section presents the used methods and data for the empirical testing of the model, while the results of the applied model are described and discussed in section four. Finally, the conclusions and recommendations for future research are given in section five.

2 Literature Review

RFM method [3] is one of the most used methods for customer value analysis. Additionally, this method is often used in direct marketing, due to its simplicity, as well as in customer purchases analysis in general. Recency represents the time since last purchase, Frequency is the number of purchases made by the customer and Monetary is a total value of purchases in the observed period. Hughes [3] approached the analysis by giving all three variables equal importance, which followed in [4]. However, a number of authors state that the importance of RFM variables should be assessed individually, which would result in attaching unequal weights to each one [5, 6]. The process of determining the weights of each variable ranges from a simple managerial judgment to its combination with analytic hierarchy process (AHP) method [7–9]. The weights give different importance to each of the RFM variables, meaning that they do not have the same influence on customer value. The RFM method has previously been used in customer value and profitability analysis in [10–14].

Liu and Shih [7] applied the AHP method to obtain the weights of RFM variables for customer lifetime value analysis by using the judgment from three groups of evaluators: administrative and business managers, marketing consultants and customers.

Based on their assessment, the relative weights of RFM variables were 0.731, 0.188 and 0.081, respectively, with Recency as the most important variable. A group of decision makers also determined the importance of each of the RFM variables for customer value categorization in [9]: manager, supervisor and administrative worker, which was an input for AHP method to determine the weights: 0.058, 0.546 and 0.395, respectively. In this case, Frequency obtained the highest weight. In [13], also, AHP, based on the view from an expert in the sales department, determined the highest value weight for the Frequency variable: 0.637, followed by Monetary with 0.258 and finally, Recency with 0.105. These weights were later used to calculate the customer lifetime value in order to later create different marketing strategies for each defined segment. Similarly, an expert group evaluated the weights of RFM variables in [15], where weights were set to: 0.17, 0.47 and 0.35, respectively, as well as in [16]: 0.3, 0.6 and 0.1, which was completely judgment-based and, in both cases, Frequency was determined to be of highest importance. This information was used to calculate RFM attributes and identify high-value and mass customers and provide insight into their preferences and patterns. Table 1 presents the RFM weights obtained in the previous studies.

In this paper, the influence of RFM attributes, as a group, as well as individually, will be assessed in predicting the customer profitability using SVR. Besides RFM attributes, depending on the availability of different data categories, product description variables, Web metrics and customer data will also be analyzed in terms of affecting the accuracy of customer profitability prediction. This approach does not rely on the judgment and potentially subjective estimations of different variables, but objectively shows, by assessing the model performance, which variables are dominant in predicting customer profitability.

Table 1 RFM weights obtained in previous research

Author(s)	Recency weight	Frequency weight	Monetary weight
Liu and Shih [7]	0.731	0.188	0.081
Monalisa et al. [9]	0.058	0.546	0.395
Khajvand et al. [13]	0.105	0.637	0.258
Safari et al. [15]	0.17	0.47	0.35
Chen and Wang [17]	0.3	0.6	0.1

3 Data and Methodology

3.1 Data

Two datasets were used for the empirical testing of the proposed approach. The first dataset is obtained from sports retailer e-commerce. This dataset is a result of merging several databases: company's own customer and product description database, as well as data from Google Analytics and Facebook Business Manager. The dataset shows the purchases of sporting goods in the period from October 2018 until January 2019 that were completed directly through targeted social media advertisements. Namely, the only traffic source in this database is the one referred from the sponsored posts on social media. Therefore, the analyzed dataset shows precisely those customers that have completed a transaction by clicking on the targeted (sponsored) Facebook or Instagram post. In the observed period, there were six direct marketing campaigns placed on social media by the company.

The final dataset was preprocessed and transformed into an appropriate form for a customer profitability prediction and split into training and test datasets of approximately the same size. The training set contains the history of Web and shopping behavior of 130 respondents (customers) from campaigns C1 to C4 and an indicator of their profitability was the value of all transactions ending with Campaign 5. The test set contains the history of web and shopping behavior of 158 respondents from campaigns C1 to C5 and the profitability indicator was the value of all transactions from the beginning of observed period, ending with Campaign 6. Hence, the lifetime profitability is being predicted, while the prediction horizon is the next campaign.

This dataset contains web log data, customer data, product data and purchasing behavior data (RFM). Customer data contains region and gender. Web metrics include: average number of sessions, average session duration, average bounce rate, number of sessions from the Central, Southern and Northern region, number of sessions using desktop, tablet and mobile devices, number of sessions using Android, iOS and Windows operating systems. Product data contains: number of purchased products in each product and brand category. RFM attributes include: Recency obtained by splitting the dataset into five equal parts from least to most recent transactions, Recency obtained by assigning numbers from 2 to 5 based on the last campaign the customer ordered in, number of campaigns with orders, total number of orders in all campaigns, number of orders in the last campaign, the average transaction amount in all campaigns, average amount of transactions in the last campaign and a total sum of realized transactions (dependent variable).

The second dataset used for model validation is the Direct Marketing Educational Foundation 3 (DMEF3) dataset, which presents catalog sales for a 12-year period. Following the procedure of data preprocessing given in [18], the present moment has been defined as August 1, 1990, and also, Winsorisation (1%) and the square root transformation to all amount and count variables were applied. The customer sample before this date was 41,669, which determined the base and target periods of

six years. In the last data preparation step, the sample was divided into training and test datasets of approximately the same size.

This dataset contains the following customer attributes: customer ID, day and year of entry in the database, time on file; RFM attributes: number of months since last order, sales amounts and number of orders by product classes in the base period, total number of orders and sales amount in the base period, whether there were orders during each observed year, whether the customer ordered in two or 3 years in a row, number of years with orders, dummy recency variables formed based on the number of months since the last order and recency quantiles 1–20; as well as lifetime attribute: lifetime sales (dependent variable). The prediction horizon in this dataset is six years.

Even though this approach in customer profitability prediction is tested on two direct marketing datasets, it can be applied in different other industries, with different available attributes.

3.2 *Methods*

The main challenge in predicting customer profitability is asymmetry since the distribution of profitability across customers is usually skewed. This problem occurs because the number of highly profitable customers is usually significantly lower compared to other, less profitable groups. In traditional regression models [6–9], the focus is on the average customer, while customer heterogeneity is usually not taken into account. In this way, not sufficiently precise distinctions are made between valuable and less valuable customers.

In line with the stated shortcoming of standard regression methods, some previous research [10, 11] has shown that these methods can hardly successfully cope with asymmetric distribution and give accurate predictions, especially for a group of very highly profitable customers. This issue is highlighted even more if customer segmentation or clustering has not been performed previously.

To estimate customer profitability, SVR method [19] was used. SVR considers data to be vectors in n -dimensional space (input space) and in case the relationship between the regressor and the dependent variable is nonlinear, the vectors are mapped to a feature space. In the feature space, it is possible to find a hyperplane that can linearly model this connection. SVR method has two aims: first, to minimize the error in the evaluation of the dependent variable, and second, to make the model in the larger dimension as flat as possible in order to increase its prediction accuracy on an unknown dataset.

Since the profitability distribution of customers is usually skewed, the accuracy of predictions decreases as the value of customer profitability increases. When the sample heterogeneity is present due to outliers, data mining methods tend to overestimate predicted amounts. In some previous research, such customers with extreme profitability values (outliers) were removed [20–22] in order to improve the prediction accuracy. However, data on such customers contains significant information for

the highly profitable customer segment, which is of great importance for decision makers. In this regard, SVR method will be used in order to overcome the problem of asymmetric variable distribution, as it was shown in some previous research [23]. Additionally, SVR works well with smaller sample sizes [24] and in this research, it will be tested on both small and large dataset.

The methodological procedure consists of the following steps:

1. Database preparation—data preprocessing, calculation of RFM attributes and normalization (range 0–1)
2. Training the SVR model for predicting the customer profitability. This step includes finding the optimal parameters for the SVR model (C , γ and ϵ), using the fivefold cross-validation and grid-search technique.
3. After the selection of the optimal parameters, the model was generated using the obtained parameter values and applied on the unseen data (test set).
4. The previous step was repeated by excluding the variable groups individually in order to assess their impact on the prediction accuracy, i.e., the importance of excluded attributes for the profitability prediction.

The same approach was applied for both datasets. The proposed predictive procedure was implemented using RapidMiner.

4 Results and Discussion

This section will provide an overview of obtained results and discussion of the proposed approach findings.

In Table 2, results on both datasets are presented. For each dataset, support vector regression was applied on all attributes, as well as on a selection of attributes, by excluding certain attributes or attribute groups. In this way, the influence on the model performance will be tested for web attributes (if applicable), RFM attributes as a group, as well as individually, product attributes and customer attributes.

By looking at the results for a sport retailer dataset, it can be seen that the model obtained a relative error (RE) of 9.19% on all data, with squared correlation (R^2) of 0.899, which indicated an overall good model performance. Taking into account the excluded attribute groups, it can be seen that the worst model performance was achieved when RFM attributes as a group were left out, which increased the RE to 29.46% and lowered the R^2 to 0.709. Another significant change in model performance occurred with Monetary variable being left out—RE grew to 31.72% and R^2 was reduced to 0.737. On the other hand, excluding web attributes and customer data has just slightly impacted the model performance with RE of 9.45 and 9.63%, and R^2 of 0.851 and 0.859, respectively. Finally, product description data being left out resulted in an improved RE of 8.29% and a lower R^2 of 0.791.

On the DMEF3 dataset, model testing on all data achieved a relative error of 25.25% and squared correlation of 0.473. The higher error level is expected due to a much longer prediction horizon (six years). Authors in [18] used this dataset

Table 2 Customer profitability prediction model results

Dataset	Performance		Training set		Test set		Customer data excluded						
			All attributes (Cross-validation)	All attributes	Web attributes excluded	Product attributes excluded	RFM attributes excluded	Recency excluded	Frequency excluded	Monetary excluded	Customer data excluded		
Sport retailer e-commerce dataset	RMSE		13.759 ± 7.634	18.246	22.238	25.968	30.541	18.843	23.826	28.580	21.629		
	Absolute error		7.315 ± 4.129	7.528 ± 16.620	8.412 ± 20.586	8.625 ± 24.494	16.427 ± 25.748	7.456 ± 17.306	10.437 ± 21.419	16.628 ± 23.245	8.482 ± 19.896		
	Relative error		9.88% ± 2.49%	9.19% ± 14.40%	9.45% ± 16.02%	8.29% ± 20.02%	29.46% ± 28.87%	8.75% ± 14.51%	13.09% ± 16.77%	31.72% ± 32.44%	9.63% ± 15.48%		
	Correlation		0.938 ± 0.040	0.948	0.922	0.889	0.842	0.944	0.908	0.858	0.927		
	Squared correlation		0.881 ± 0.075	0.899	0.851	0.791	0.709	0.892	0.825	0.737	0.859		
DMEF3 dataset	Spearman rho		0.938 ± 0.047	0.937	0.925	0.899	0.737	0.94	0.916	0.715	0.929		
	RMSE		5.407 ± 0.081	5.334	-	5.360	6.952	5.359	5.517	5.581	5.721		
	Absolute error		4.024 ± 0.043	3.995 ± 3.535	-	4.029 ± 3.536	5.183 ± 4.633	4.015 ± 3.550	4.131 ± 3.656	4.181 ± 3.697	4.308 ± 3.765		
	Relative error		25.50% ± 0.52%	25.25% ± 21.84%	-	25.58% ± 22.16%	32.77% ± 28.95%	25.39% ± 21.90%	26.13% ± 22.68%	26.45% ± 23.07%	27.18% ± 22.46%		
	Correlation		0.691 ± 0.014	0.688	-	0.684	0.340	0.685	0.661	0.650	0.632		
Spearman rho	Squared correlation		0.478 ± 0.019	0.473	-	0.468	0.116	0.469	0.437	0.422	0.399		
	Spearman rho		0.478 ± 0.019	0.647	-	0.638	0.360	0.642	0.622	0.621	0.555		

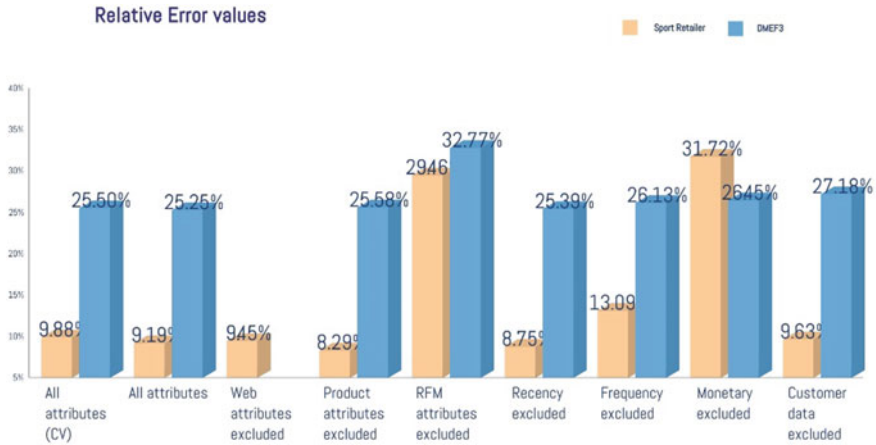


Fig. 1 Relative error (RE) values across tested models for both datasets

for profitability prediction using linear regression and achieved the misclassification rate of 55% of the top 20% customers. Authors discussed whether it is possible to accurately predict the profitability in such a long prediction horizon. However, the aim of this study is to assess the prediction accuracy in relation to the attributes used.

Similar to the first dataset, the left out RFM variables lead to significant worsening of the model performance with RE and R^2 equal to 32.77% and 0.116, respectively. However, customer data had more impact here, as the model reached second worst performance levels (RE = 27.18% and $R^2 = 0.399$). Slight changes occurred by leaving out product description data, as well as RFM variables individually. The largest impact is obtained by leaving out the Monetary variable, followed by Frequency and Recency, respectively. However, the differences are smaller compared to the first dataset, due to a longer prediction horizon. Relative error and squared correlation values obtained in all tested models are presented in Figs. 1 and 2.

Overall, the results show that adequate prediction can be made using RFM attributes, if that is all that is available to a company. RFM attributes contain significant information about customers' purchasing history and in both tested datasets, these variables proved to be the most significant predictor of customer profitability. This is in line with results from previous research, which highlights the importance of RFM attributes for customer value estimation and profitability prediction [25, 26].

By analyzing these attributes individually, they would be ranked in the following order: M, F and R. In the case of both datasets, Monetary variable determines the customer profitability the most. This is not in line with most previous research surveyed in literature review section. In [9, 13, 15], Monetary variable was second in rank of importance, while in [7, 16], it was the least important of the three. However, the fuzzy-AHP technique in [27] reached the same order of importance of RFM variables, as in this study. Hence, the importance of RFM variables as a group is undoubtable; however, their individual impact on customer profitability should be

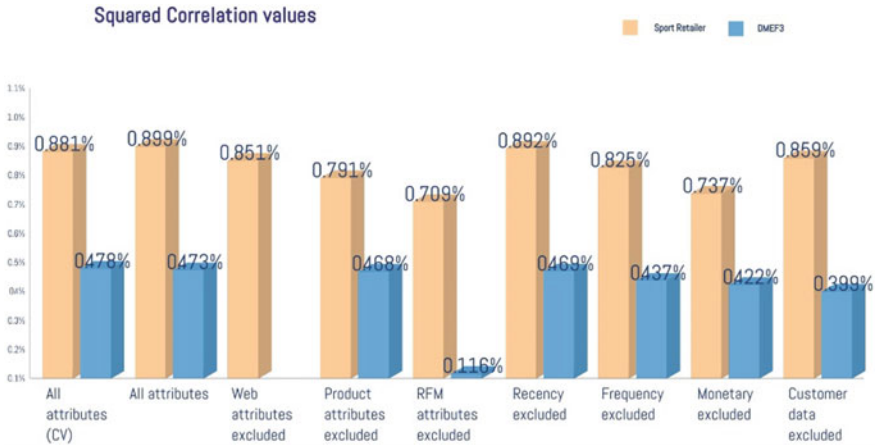


Fig. 2 Squared correlation (SC) values across tested models for both datasets

explored in detail in each individual case, industry or sector, as mentioned in [28]. Even though our results suggest that monetary variable has the most significant effect on the profitability prediction, with longer prediction horizon, its leading position reduces and all three RFM attributes have a similar importance in that case. Hence, the larger influence of the Monetary variable in the first dataset can be explained by a shorter prediction horizon—if a customer had substantial transactions in the previous campaigns, these amounts will be transferred to the following period’s sum of transactions. On the other hand, if the case was to predict for the longer term (several years, for instance), the transaction amounts from the first part of the observed period would not dictate the ultimate profitability in the long run.

When it comes to web metrics, this research did not prove their significance for predicting customer value. In previous research, authors pointed out that web log data represent predominant factors for customer segmentation, based on their history of purchases [29]. Additionally, web metrics proved significant in customer response modeling and predicting future purchases [30, 31]. Therefore, web log data and metrics can be used in several other customer analyses, but in the case of prediction profitability, this attribute group did not prove to be significant. In that sense, if average session duration, for example, is used in customer response modeling, it could be expected to affect this prediction—it can determine whether the transaction will occur or not, but it is less likely that it will determine the transaction value itself. Usually, when evaluating the significance of attributes, managers have in mind the customer’s response and make judgments based on that. However, when it comes to profitability prediction based on the transaction value, then it is more important whether the customer spent a lot or a little in the past. Hence, our results suggest that for the profitability prediction, in comparison to response modeling, focus should be put on the RFM, and especially Monetary attribute in the short term.

Utilizing different categories of data can help in profiling and segmenting the customers, predicting their profitability and creating personalized messages and offers for their value maximization.

5 Conclusion

In this paper, the significance of RFM attributes for the prediction of customer profitability was tested, in relation to web metrics, customer and product description data. For the prediction of profitability, as a continuous variable, SVR approach was implemented. This study uniquely uses several different data categories and sources and potentially new performance prediction factors that prior research has rarely investigated.

The results of this study point out that companies could rely on RFM attributes exclusively to analyze and predict the profitability of their customers. However, purchasing behavior data can be used in combination with other data categories, such as web metrics or customer data, but this data type proved its value for customer profitability prediction, which is also confirmed in previous literature. This study highlighted the Monetary attribute for a short-term prediction of customer profitability. In the case of sport retailer dataset, excluding the Monetary attribute led to an increase of the relative error to the value of 31.72%, which was the highest level of error compared to all other tested models. Additionally, the results point to the decreasing significance of this attribute with the increasing prediction horizon. In longer term predictions, all three RFM attributes manifested similar effect on the prediction performance. When it comes to the Monetary attribute and its impact in the DMEF3 dataset, the RE levels across all RFM attributes are similar (26.13%, 26.45% and 27.18%, respectively) which highlights the obtained conclusion.

From a marketing standpoint, presented results can be a foundation for customer segmentation and target selection based on profitability, for future direct marketing activities. As a recommendation for future research, a combination of customer profitability prediction, and customer response modeling can be conducted to obtain a full picture of most profitable customers as well as those with highest probability of a response. Information from such research can be particularly useful in companies dealing with marketing budget constraints.

References

1. Palmer, A., Koenig-Lewis, N.: An experiential, social network-based approach to direct marketing. *Direct Mark. Int. J.* **3**, 162–176 (2009). <https://doi.org/10.1108/17505930910985116>
2. Sheikh, A., Ghanbarpour, T., Gholamiangonabadi, D.: A preliminary study of Fintech industry: a two-stage clustering analysis for customer segmentation in the B2B setting. *J. Bus. Mark.* **26**, 197–207 (2019). <https://doi.org/10.1080/1051712X.2019.1603420>

3. Hughes, A.M.: *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*. Irwin, Chicago (1994)
4. Cheng, C.H., Chen, Y.S.: Classifying the segmentation of customer value via RFM model and RS theory. *Expert Syst. Appl.* **36**, 4176–4184 (2009). <https://doi.org/10.1016/j.eswa.2008.04.003>
5. McCarty, J.A., Hastak, M.: Segmentation approaches in data-mining: a comparison of RFM, CHAID, and logistic regression. *J. Bus. Res.* **60**, 656–662 (2007). <https://doi.org/10.1016/j.jbusres.2006.06.015>
6. Stone, B.: *Successful Direct Marketing Methods*. NTC Business Books, Lincolnwood, IL (1995)
7. Liu, D.R., Shih, Y.Y.: Integrating AHP and data mining for product recommendation based on customer lifetime value. *Inf. Manage.* **42**, 387–400 (2005)
8. Yao, L., Xiong, J.: Customers segmentation using RFM and two-step clustering. *Adv. Mater. Res.* **268–270**, 631–635 (2011)
9. Monalisa, S., Nadya, P., Novita, R.: Analysis for customer lifetime value categorization with RFM model. *Procedia Comput. Sci.* **161**, 834–840 (2019). <https://doi.org/10.1016/j.procs.2019.11.190>
10. Heldt, R., Silveira, C.S., Luce, F.B.: Predicting customer value per product: from RFM to RFM/P. *J. Bus. Res.* (2019). <https://doi.org/10.1016/j.jbusres.2019.05.001>
11. Rogic, S., Kascelan, L.: Customer value prediction in direct marketing using hybrid support vector machine rule extraction method. *Commun. Comput. Inf. Sci.* **1064**, 283–294 (2019). https://doi.org/10.1007/978-3-030-30278-8_30
12. Doğan, O., Aycin, E., Bulut, Z.A.: Customer segmentation by using RFM model and clustering methods: a case study in retail industry. *Int. J. Contemp. Econ. Adm. Sci.* **8**, 1–19 (2018)
13. Khajvand, M., Zolfaghar, K., Ashoori, S., Alizadeh, S.: Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Comput. Sci.* **3**, 57–63 (2011). <https://doi.org/10.1016/j.procs.2010.12.011>
14. Chen, D., Guo, K., Li, B.: Predicting customer profitability dynamically over time: an experimental comparative study. In: *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes Bioinformatics)*, vol. 11896 LNCS, pp. 174–183 (2019)
15. Safari, F., Safari, N., Montazer, G.A.: Customer lifetime value determination based on RFM model. *Mark. Intell. Plan.* **34**, 446–461 (2016). <https://doi.org/10.1108/MIP-03-2015-0060>
16. Chen, Q., Zhang, M., Zhao, X.: Analysing customer behaviour in mobile app usage. *Ind. Manag. Data Syst.* **117**, 425–438 (2017). <https://doi.org/10.1108/IMDS-04-2016-0141>
17. Chen, K.Y., Wang, C.H.: Support vector regression with genetic algorithms in forecasting tourism demand. *Tour. Manag.* **28**, 215–226 (2007)
18. Malthouse, E.C., Blattberg, R.C.: Can we predict customer lifetime value? *J. Interact. Mark.* **19**, 2–16 (2005). <https://doi.org/10.1002/dir.20027>
19. Vapnik, V., Golowich, S.E., Smola, A.: Support vector method for function approximation, regression estimation, and signal processing. *Adv. Neural Inf. Process. Syst.* 281–287 (1997)
20. Lei, M., Jiang, G., Yang, J., Mei, X., Xia, P., Shi, H.: Improvement of the regression model for spindle thermal elongation by a Boosting-based outliers detection approach. *Int. J. Adv. Manuf. Technol.* **99**, 1389–1403 (2018). <https://doi.org/10.1007/s00170-018-2559-8>
21. Wang, K., Lan, H.: Robust support vector data description for novelty detection with contaminated data. *Eng. Appl. Artif. Intell.* **91**, 103554 (2020)
22. Kim, D., Lee, H. joo, Cho, S.: Response modeling with support vector regression. *Expert Syst. Appl.* **34**, 1102–1108 (2008). <https://doi.org/10.1016/j.eswa.2006.12.019>
23. Kaščelan, V., Kaščelan, L., Burić, M.N.: A nonparametric data mining approach for risk prediction in car insurance: a case study from the Montenegrin market. *Econ. Res. Istraz.* **29**, 545–558 (2016). <https://doi.org/10.1080/1331677X.2016.1175729>
24. Tange, R.I., Rasmussen, M.A., Taira, E., Bro, R.: Benchmarking support vector regression against partial least squares regression and artificial neural network: effect of sample size on model performance. *J. Near Infrared Spectrosc.* **25**, 381–390 (2017)

25. Wei-jiang, L., Shu-yong, D., Xue, Y., Xiao-feng, W.: Determination of customer value measurement model RFM index weights. *African J. Bus. Manage.* **5**, 5567–5572 (2011)
26. Ait Daoud, R., Bouikhalene, B., Amine, A., Lbibb, R.: Combining RFM model and clustering techniques for customer value analysis of a company selling online. In: 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA). IEEE (2015)
27. Aggarwal, A.G., Delhi, N.: Customer segmentation using fuzzy-AHP and RFM model. In: 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 77–80 (2020)
28. Dursun, A., Caber, M.: Using data mining techniques for profiling prof table hotel customers: an application of RFM analysis. *Tour. Manage. Perspect.* **18**, 153–160 (2016)
29. Van den Poel, D., Buckinx, W.: Predicting online-purchasing behaviour. *Eur. J. Oper. Res.* **166**, 557–575 (2005)
30. Esmeli, R., Bader-El-Den, M., Abdullahi, H.: Towards early purchase intention prediction in online session based retailing systems. *Electron. Mark.* (2020)
31. Martínez, A., Schmuck, C., Pereverzyev, S., Pirker, C., Haltmeier, M.: A machine learning framework for customer purchase prediction in the non-contractual setting. *Eur. J. Oper. Res.* **281**, 588–596 (2020)

Marketing, Geomarketing and IOT

Toward Consumer Sustainable Consumption: Examining Factors Influencing Green Product Purchase Intention



Marleen Prigita and Yeshika Alversia

Abstract The rapid growth of the industrial sector and population causes waste problems in Indonesia. The government began implementing regulations to limit the use of plastics. However, this effort needs support from both consumers and companies. Consumers are expected to adopt sustainable consumption in their daily life, and companies are urged to involve environmental sustainability in their strategy. Thus, green products appear as one of the possible solutions for reducing environmental sustainability issues. This study aims to analyze factors that could influence green product purchase intention for generation Y and generation Z in Indonesia. Generation Y and generation Z were known for their concerns toward the environment, and they are potential current and future consumers. Therefore, it is essential to understand their consumption behavior. This study is based on the theory of planned behavior framework that was extended with additional variables, namely environmental concern, environmental knowledge, willingness to pay premium, moral attitude, and health consciousness. Data were collected by distributing questionnaires to 349 respondents adopting purposive sampling approach. Data were analyzed using structural equation modeling. The result reported that all variables, except environmental concern, have positive influence on purchase intention.

Keywords Green product · Purchase intention · Sustainability

1 Introduction

To reduce environmental impact, the Indonesian government begins implementing regulations which are consistent with sustainable development goals by the United Nations. However, these efforts need support from both consumers and companies by changing their consumption and production pattern.

M. Prigita (✉) · Y. Alversia
Universitas Indonesia, Jl. Salemba Raya No. 4, Jakarta 10430, Indonesia
e-mail: marleen.prigita@ui.ac.id

Y. Alversia
e-mail: yeshika@ui.ac.id

For consumers, the first step to protect the environment can be taken by implementing sustainable consumption in their everyday life. Sustainable consumption refers to products or services consumption that responds to meet basic necessities and improve quality of life while reducing the use of natural resources and toxic materials in order not to hinder future generation's ability to fulfill their needs [1].

On the other hand, companies are urged to involve environmental sustainability in their strategies [2] as an embodiment of the triple bottom line concept. Triple bottom line is the basis of a sustainable strategy and will help companies achieve sustainable competitive advantage.

Thus, green products appear as a possible solution for achieving sustainable production and consumption, which will further reduce environmental sustainability issues. Green products refer to products that can minimize environmental damage or products that are designed using recycled materials [3].

Millennials (generation *Y*) are the largest working-age population groups that play an essential role in Indonesia's economy [4]. Millennials in Indonesia were having a serious concern for the plastic waste problem and trying to adopt an environmentally friendly lifestyle. They prefer brand and product campaigns that positively contribute to the environment and surrounding communities [5].

Meanwhile, generation *Z* is a generation group that was born between 1995 and 2010. Thus, generation *Z* also contributes to the economy of Indonesia. Generation *Z* is known to be sensitive and concerned about the environment [6].

In Indonesia, prior research had been conducted on green product purchase intention. However, there is still very limited research using generation *Y* and generation *Z* respondents. As both generations are potential current and future consumers, their standpoint represents the community [7]. Thus, it is necessary to understand their consumption behavior. Moreover, no previous studies have developed the TPB framework by adding variables: environmental concern, environmental knowledge, health consciousness, moral attitude, and willingness to pay premium. In fact, these variables can influence consumer purchase intention. Thus, this study aims to understand the factors and the determinants of green product purchase intention for generation *Y* and generation *Z* in Indonesia.

2 Literature Review

2.1 Environmental Concern (EC)

Consumers with a greater environmental concern will tend to protect the environment and are willing to alter their consumption patterns to reduce environmental impact [8]. Thus, they tend to show a greater chance of buying green products [9].

H_1 : Environmental concern positively influences attitude toward buying green product.

H_2 : Environmental concern positively influences purchase intention green product.

2.2 Environmental Knowledge (EK)

Environmental knowledge refers to knowledge of facts and concepts about the environment [10]. Information about consumers' environmental knowledge is useful for evaluating green consumer behavior and developing strategies [11].

H_3 : Environmental knowledge positively influences attitude toward buying green product.

H_4 : Environmental knowledge positively influences purchase intention green product.

2.3 Theory of Planned Behavior

The theory of planned behavior has three variables: subjective norm, attitude, and perceived behavioral control. Attitude (ATT) is a psychological emotion that is diverted through consumer evaluation, and if positive, then behavioral intention will tend to be more positive [12]. Subjective norm (SN) is a person's belief that is influenced by other people to perform certain behavior [13]. Under the influence of subjective norms, an individual can feel the social pressure to participate in eco-friendly activities, such as buying green products to match their role model's expectations. Perceived behavioral control (PBC) is an easy or difficult feeling to do a behavior and reflects past experience and anticipation of obstacles that are expected to emerge [14].

H_5 : Attitude toward buying green product positively influences purchase intention green product.

H_6 : Subjective norm positively influences purchase intention green product.

H_7 : Perceived behavioral control positively influences purchase intention green product.

2.4 Willingness to Pay Premium (WTP)

Green products generally have a higher selling price than conventional products because they require higher costs in the production process [15]. Therefore, the premium price can be a barrier to green product consumption, especially for young people with limited purchase abilities [16]. Willingness to pay premium means the maximum price can be accepted by a consumer to pay for a product or service [17].

H_8 : Willingness to pay premium positively influences purchase intention green product.

2.5 Moral Attitude (MA)

To overcome the theory of planned behavior's limitation of moral influence on behavior, Arvola et al. introduced moral attitude, which is a moral norm that is operationalized as positive self-enhancing feelings of doing the right thing, rather than negative feelings of obligation or guilt [18].

H_9 : Moral attitude positively influences purchase intention green product.

2.6 Health Consciousness (HC)

Health consciousness indicates the degree to which a person participates in health-related actions and therefore reflects health-oriented choices [19].

H_{10} : Health consciousness positively influences attitude toward buying green product.

H_{11} : Health consciousness positively influences purchase intention green product.

2.7 Purchase Intention (PI)

Green purchase intention is defined as the extent to which a consumer is ready to buy eco-friendly products which have a low impact on the environment [20]. Various studies reported determinants of green purchase intention, e.g., eco-concerns [21], knowledge [7], attitude [22].

The conceptual model was developed by combining three previous studies [21–23], which all used the basic framework of theory of reasoned action (TRA) and theory of planned behavior (TPB), then extended by adding new latent variables. The proposed conceptual model has nine latent variables, as shown in Fig. 1.

3 Methodology

Based on the preliminary survey result, green products discussed in this study were limited in the beauty and personal care category. The primary data were collected from a survey by distributing the questionnaire. Respondents are required to answer 44 questions in the questionnaire. Attitude toward buying green product was measured on a six points semantic differential scale, while the remaining latent variables were measured with six points Likert scale, where (1) denotes strong disagreement and (6) denotes strong agreement.

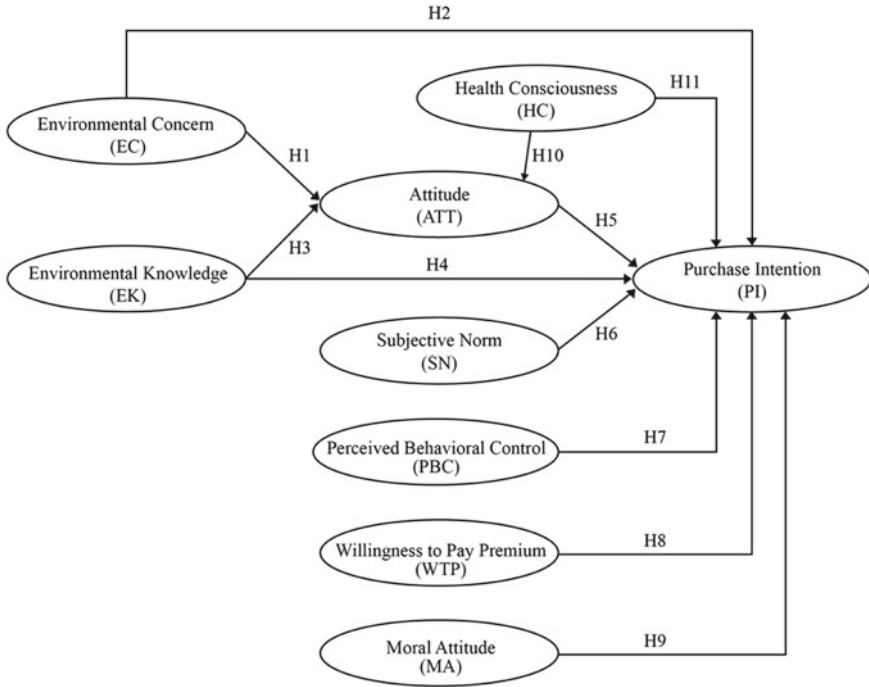


Fig. 1 Conceptual model

Data were further analyzed using structural equation modeling (SEM). The minimum amount of sample needed in this study was five respondents per indicator [24]. There were 349 respondents participated in this survey which was higher than the recommended amount of 220 respondents ($= 5 \times 44$ indicator). This study adopted a purposive sampling technique for collecting respondents. Respondents of this study were generation Z (17–25 years old) and generation Y (26–40 years old) who live in the Greater Jakarta area.

4 Results and Discussion

4.1 Respondent’s Profile

Most participants were female (63.9%), aged between 26 and 30 years old (33.8%), hold a bachelor’s degree (61.3%), worked as an employee (50.1%), and lived in Jakarta (64.1%) with monthly expenditure between R_p 2,000,001 and R_p 4,000,000 (32.7%).

4.2 Measurement Model Analysis

The *t*-value of indicators ranged from 15.89 to 22.04. Standardized factor loading value ranged from 0.75 to 0.91. The *t*-value of all indicators ≥ 1.65 , and standardized factor loading value ≥ 0.70 , then the validity requirements were met.

Meanwhile, the construct reliability value of all variables ranged from 0.865 to 0.942. The variance extracted of all variables ranged from 0.631 to 0.766. The construct reliability value was higher than 0.70, and the variance extracted value was higher than 0.5 indicate that all latent variables were reliable. In conclusion, all variables and indicators were valid and reliable to be used for further testing.

Next, the goodness of fit (GOF) of the measurement model was tested. GFI = 0.80 (marginal fit), while RMR = 0.046, RMSEA = 0.06, NNFI = 0.99, NFI = 0.98, RFI = 0.98, IFI = 0.99, and CFI = 0.99 (good fit). Overall, the measurement model was identified as good fit because the results obtained ten good fit and one marginal fit.

4.3 Structural Model Analysis

In this stage, the GOF of the structural model was examined. GFI = 0.80 (marginal fit), while RMR = 0.049, RMSEA = 0.06, NNFI = 0.99, NFI = 0.98, RFI = 0.98, IFI = 0.99, and CFI = 0.99 (good fit). Overall, the structural model was identified as good fit because it obtained ten good fit and one marginal fit result.

As shown in Table 1, of the eleven hypotheses proposed, ten hypotheses were accepted (indicated a significant relationship) because the *t*-value was above 1.65. Meanwhile, hypothesis H_2 was rejected because the *t*-value = 1.08, which was below 1.65. All the standardized coefficients estimates (β -value) were positive indicate

Table 1 Hypothesis testing result

Path	Estimate	<i>t</i> -value	Conclusion	Previous research
H_1 : EC \rightarrow ATT	0.61	11.02	Accepted	[12, 23]
H_2 : EC \rightarrow PI	0.05	1.08	Rejected	[21]
H_3 : EK \rightarrow ATT	0.20	3.48	Accepted	[23, 25]
H_4 : EK \rightarrow PI	0.12	2.62	Accepted	[23, 26]
H_5 : ATT \rightarrow PI	0.16	3.05	Accepted	[23]
H_6 : SN \rightarrow PI	0.12	2.49	Accepted	[23]
H_7 : PBC \rightarrow PI	0.17	2.68	Accepted	[23]
H_8 : WTP \rightarrow PI	0.19	4.24	Accepted	[22]
H_9 : MA \rightarrow PI	0.17	2.33	Accepted	[21, 27]
H_{10} : HC \rightarrow ATT	0.21	3.77	Accepted	[21, 28]
H_{11} : HC \rightarrow PI	0.14	2.76	Accepted	[21]

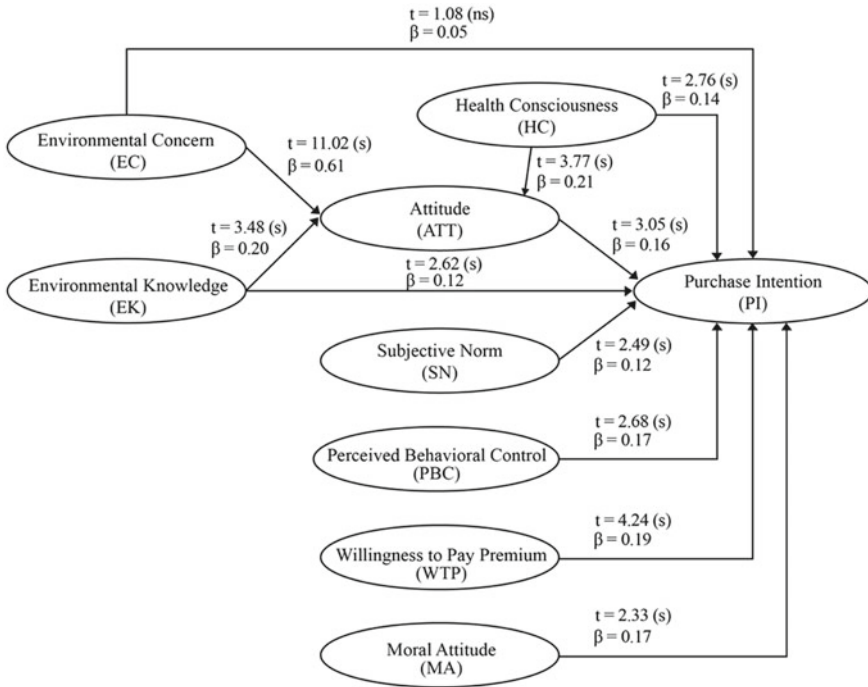


Fig. 2 Structural model

that the causal relationships formed were positive. The relationship between latent variables can be seen in Fig. 2.

4.4 Discussion

As shown in Fig. 2, environmental concern significantly influenced attitude toward buying green product and supported hypothesis 1. This indicates, the greater someone’s concern for environmental issues, he will show an increasingly positive attitude toward buying green product.

Environmental concern did not significantly influence on purchase intention because the t -value = 1.08, which was smaller than the t -table of 1.65 at a 95% confidence level. Thus, hypothesis 2 was rejected. A similar result was found in the previous research [21]. The reasons consumers purchase organic products can be grouped into two categories, namely egoistic reasons and altruistic reasons [29]. The environmental concern motive is more altruistic because it prioritizes the interests of the community rather than personal interests [30]. Most consumers are unwilling to give up their personal interests in order to benefit the community [29]. Therefore, in making green product decisions, egoistic motives are considered more important

than altruistic motives such as environmental concern [30], and this can be one of the causes of the insignificant relationship between environmental concern and purchase intention.

Environmental knowledge significantly influenced attitude toward buying green product and supported hypothesis 3. This indicates, the greater someone's environmental knowledge, he will show an increasingly positive attitude toward buying green product.

Environmental knowledge significantly influenced purchase intention and supported hypothesis 4. This indicates, the greater someone's environmental knowledge, he will show a greater intention to purchase green product.

Attitude toward buying green product significantly influenced purchase intention and supported hypothesis 5. This indicates, the more positive someone's attitude toward buying green product, he will show a greater intention to purchase green product [31].

Subjective norm significantly influenced purchase intention and supported hypothesis 6. This indicates, the greater the subjective norm, the greater the purchase intention. In other words, the greater influence of meaningful people around someone to buy green product, then he will tend to show a greater intention to purchase green product.

Perceived behavioral control significantly influenced purchase intention and supported hypothesis 7. This indicates, the more a person is able to control factors that can complicate his behavior of buying green products, he will show greater intention to purchase green product.

Willingness to pay premium significantly influenced purchase intention and supported hypothesis 8. This indicates, the higher the willingness to pay premium, the higher the intention to buy green product. That is, it requires willingness to pay premium for generation *Y* and generation *Z* to buy green product.

Moral attitude significantly influenced purchase intention and supported hypothesis 9. This indicates someone with a greater moral attitude will have more intention to purchase green product.

Health consciousness significantly influenced attitude toward buying green product and supported hypothesis 10. This indicates someone with a higher health consciousness tends to show a more positive attitude toward buying green product.

Health consciousness significantly influenced purchase intention and supported hypothesis 11. This indicates, the higher someone's health consciousness, the greater the intention to purchase green product.

There were several indirect relationships in the model. First, the indirect effect of environmental concern to purchase intention via attitude toward green product was $0.61 \times 0.16 = 0.0976$. Second, the indirect effect of environmental knowledge to purchase intention was $0.20 \times 0.16 = 0.032$. Third, the indirect effect of health consciousness to purchase intention was $0.21 \times 0.16 = 0.0336$.

Meanwhile, the total effect from environmental concern to purchase intention was $0.05 + 0.0976 = 0.1476$, environmental knowledge to purchase intention was $0.12 + 0.032 = 0.152$, and health consciousness to purchase intention was $0.14 + 0.0336 = 0.1736$.

Table 2 Result of comparison between generation Y and generation Z

Variable	Sig. value Levene's test	Conclusion	Value of equality of means		Conclusion
			Equal variances assumed	Equal variances not assumed	
ATT	0.732	Homogeneous	0.419	–	No differences
SN	0.024	Not homogeneous	–	0.150	No differences
PBC	0.749	Homogeneous	0.588	–	No differences
EC	0.833	Homogeneous	0.822	–	No differences
EK	0.737	Homogeneous	0.372	–	No differences
WTP	0.936	Homogeneous	0.208	–	No differences
MA	0.327	Homogeneous	0.107	–	No differences
HC	0.342	Homogeneous	0.285	–	No differences
PI	0.980	Homogeneous	0.422	–	No differences

The R-squared (R_2) value of attitude toward buying green product (ATT) is 0.74, indicates that 74% of the variance of attitude toward buying green product can be explained by EC, EK, and HC. Meanwhile, the R-squared (R_2) value of purchase intention (PI) is 0.94, indicates that 94% of the variance of purchase intention can be explained by EK, ATT, SN, PBC, WTP, MA, and HC.

Next, the independent sample *t*-test was conducted to identify any perception differences between generation Y and generation Z in each latent variable by comparing means between two independent groups [32]. As shown in Table 2, in conclusion, there is no perception differences between generation Y and generation Z in each latent variable.

5 Conclusion

The findings showed that green product purchase intention was influenced by seven factors, namely subjective norm, attitude toward buying green product, perceived behavioral control, environmental knowledge, health consciousness, willingness to pay premium, and moral attitude. Meanwhile, attitude toward buying green product was strongly influenced by environmental concern, followed by health consciousness and environmental knowledge. Therefore, environmental concern was indirectly influenced purchase intention, through attitude toward buying green product. The findings provide valuable insights for marketers in developing green product strategies, such as

- As generation Y and generation Z have serious concerns regarding environmental issues, marketers should consider generation Y and generation Z as their target market. Convincing generation Y and generation Z about the benefit of green

products may provide companies additional profits as they are willing to spend more money when buying green product.

- Willingness to pay premium has the strongest influence on purchase intention. To encourage consumers to be willing to pay a premium price, marketers should communicate the value and benefit of green products honestly and responsibly so the consumers will understand the benefit they get for the premium price they pay, and they will not feel deceptive by marketing claims.
- Attitude toward buying green product, the second strongest predictor, is strongly influenced by environmental concern, followed by health consciousness and environmental knowledge. This means, marketers should develop consumers' positive attitude by encouraging consumers to be more concerned to the environment and provide knowledge about the environment through education sessions or CSR activities on health and environmental sustainability.
- Marketers are also recommended to increase consumers' health consciousness by communicating the importance of health awareness through education sessions and campaigns.
- Marketers should enhance customer convenience to find green products and develop good distribution channels to ensure green products' availability in both offline and online stores.

Future research is recommended to identify other factors that could influence consumers' intention when buying green products include other age groups, increase the number of samples, expand the geographic area of the study, use other categories of green products (e.g., food, clothing), and include the measurement of behavior.

References

1. de Moura, A. P., Cunha, L. M., Castro-Cunha, M., Lima, R. C.: A comparative evaluation of women's perceptions and importance of sustainability in fish consumption: an exploratory study among light consumers with different education levels. *Manag. Environ. Qual. Int. J.* **23**(4) (2012)
2. Unruh, G., Ettenson, R.: Growing green. *Harvard Bus. Rev.* **88**(6) (2010)
3. Durif, F., Boivin, C., Julien, C.: In search of a green product definition. *Innov. Market.* **6**(1) (2010)
4. BPS.: Statistik Gender Tematik: Profil Generasi Milenial Indonesia. Kementerian Pemberdayaan Perempuan dan Perlindungan Anak, Jakarta (2018)
5. IDN Homepage, <https://cdn.idntimes.com/content-documents/Indonesia-millennial-report-2020-by-IDN-Research-Institute.pdf>. Last Accessed 2020/12/22
6. <https://www.forbes.com/sites/emanuelabarbiroglio/2019/12/09/generation-z-fears-climate-change-more-than-anything-else/?sh=78c31867501b>. Last Accessed 2020/12/22
7. Kanchanapibul, M., Lacka, E., Wang, X., Chan, H.K.: An empirical investigation of green purchase behaviour among the young generation. *J. Cleaner Prod.* **66** (2014)
8. Tan, C.N.L., Ojo, A.O., Thurasamy, R.: Determinants of green product buying decision among young consumers in Malaysia. *Young Consum.* **20**(2) (2019)
9. Kalafatis, S.P., Pollard, M., East, R., Tsogas, M.H.: Green marketing and Ajzen's theory of planned behaviour: a cross-market evaluation. *J. Consum. Market.* **16**(5) (1999)

10. Fryxell, G.E., Lo, C.W.H.: The influence of environmental knowledge and values on managerial behaviours on behalf of the environment: an empirical examination of managers in China. *J. Bus. Ethics* **46**(1), 45–69 (2003)
11. Cheah, I., Phau, I.: Attitudes towards environmentally friendly products: the influence of ecoliteracy, interpersonal influence and value orientation. *Market. Intell. Plann.* **29**(5) (2011)
12. Chen, M.F., Tung, P.J.: Developing an extended theory of planned behavior model to predict consumers' intention to visit green hotels. *Int. J. Hosp. Manag.* **36**, 221–230 (2014)
13. Paek, H.-J., Yu, J. (Hyunjae), Bae, B.J.: Is on-line health promotion culture-bound? Cultural characteristics manifested in U.S. and South Korean antismoking web sites. *J. Advertising* **38**(1), 35–48 (2009).
14. Ajzen, I.: The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **50**(2), 179–211 (1991)
15. Dangelico, R.M., Vocalelli, D.: “Green Marketing”: aAn analysis of definitions, strategy steps, and tools through a systematic review of the literature. *J. Cleaner Prod.* **165** (2017)
16. Chaudhary, R., Bisai, S.: Factors influencing green purchase behavior of millennials in India. *Manag. Environ. Qual. Int. J.* **29**(5) (2018)
17. Le Gall-Ely, M.: Definition, measurement and determinants of the consumer's willingness to pay: a critical synthesis and avenues for further research. *Recherche et Appl. En Market. (English Edition)* **24**(2) (2009)
18. Arvola, A., Vassallo, M., Dean, M., Lampila, P., Saba, A., Lähteenmäki, L., Shepherd, R.: Predicting intentions to purchase organic food: the role of affective and moral attitudes in the theory of planned behaviour. *Appetite* **50**(2–3) (2008)
19. Moorman, C., Matulich, E.: A model of consumers' preventive health behaviors: the role of health motivation and health ability. *J. Consum. Res.* **20**(2) (1993)
20. Chen, M.F., Lee, C.L.: The impacts of green claims on coffee consumers' purchase intention. *Br. Food J.* **117**(1) (2015)
21. Yadav, R., Pathak, G.S.: Intention to purchase organic food among young consumers: evidences from a developing nation. *Appetite* **96** (2016a)
22. Prakash, G., Pathak, P.: Intention to buy eco-friendly packaged products among young consumers of India: a study on developing nation. *J. Cleaner Prod.* **141** (2017)
23. Yadav, R., Pathak, G.S.: Young consumers' intention towards buying green products in a developing nation: extending the theory of planned behavior. *J. Cleaner Prod.* **135** (2016b)
24. Ferdinand, A.: Metode Penelitian Manajemen: Pedoman Penelitian untuk Skripsi. Tesis dan Disertasi Ilmu Manajemen. Badan Penerbit Universitas Diponegoro, Semarang (2016)
25. Chan, R.Y.K.: Determinants of Chinese consumers' green purchase behavior. *Psychol. Market.* **18**(4) (2001)
26. Wang, P., Liu, Q., Qi, Y.: Factors influencing sustainable consumption behaviors: a survey of the rural residents in China. *J. Cleaner Prod.* **63** (2014)
27. Liu, M.T., Liu, Y., Mo, Z.: Moral norm is the key: an extension of the theory of planned behaviour (TPB) on Chinese consumers' green purchase intention. *Asia Pac. J. Market. Logistics* **32**(8) (2020)
28. Lee, H.J.: Individual and situational determinants of U.S. consumers' buying behavior of organic foods. *J. Int. Food Agribus. Market.* **28**(2) (2016)
29. Wandel, M., Bugge, A.: Environmental concern in consumer evaluation of food quality. *Food Qual. Preference* **8**(1) (1997)
30. Magnusson, M.K., Arvola, A., Hursti, U.K.K., Åberg, L., Sjöden, P.O.: Choice of organic foods is related to perceived consequences for human health and to environmentally friendly behaviour. *Appetite* **40**(2), 109–117 (2003)
31. Kashi, A.N.: Green purchase intention: a conceptual model of factors influencing green purchase of Iranian consumers. *J. Islamic Market.* **11**(6) (2019)
32. Sekaran, U.: *Research Methods for Business: A Skill Building Approach*. Wiley, United Kingdom (2016)

A Smart City IoT Crowdsensing System Based on Data Streaming Architecture



Aleksandra Labus , Miloš Radenković , Stefan Nešković,
Snežana Popović , and Svetlana Mitrović

Abstract The subject of this paper is data streaming in IoT crowdsensing systems. The goal of this paper is to present a way of designing a scalable IoT crowdsensing system that enables design of various business models for smart city projects. The system designed in such a way is capable of handling an increasing number of users while maintaining acceptable performance. Performance of the system can be measured in response latency, which allows for real-time tracking of crowdsensing parameters. The first part of the paper deals with data streaming concepts and software solutions, with a particular focus on Apache Kafka. The second part presents the designed system for crowdsensing in smart city environments. The designed system allows for use of mobile and Arduino devices as input data for the Kafka cloud cluster in order to provide crowdsourcing insights in real-time. The primary way that users can utilize these insights is through a web or mobile application, where various data visualizations can be presented. The development of a system based on the proposed model can allow for easy access to recent crowdsourced data, and real-time smart city indicators such as air pollution.

Keywords Data streaming · Crowdsensing · Mobile crowdsensing · IoT

A. Labus (✉) · S. Nešković

Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, 11000 Belgrade, Serbia
e-mail: aleksandra@elab.rs

S. Nešković

e-mail: stefanneskovic@elab.rs

M. Radenković · S. Popović

School of Computing, Union University, Knez Mihailova 6/VI, 11000 Belgrade, Serbia
e-mail: mradenkovic@raf.rs

S. Popović

e-mail: spopovic@raf.rs

S. Mitrović

Faculty of Project and Innovation Management, Educons University, Bože Jankovića 14, 11000 Belgrade, Serbia
e-mail: svmitrovic@mts.rs

1 Introduction

With the increasing needs of enterprises and reliance on the internet both on the consumer and enterprise side, the ways of utilizing data is quickly changing. The conventional use of data revolved around processing it, usually in a batch-processing fashion, and then storing it in a database of some kind for later analysis and user query. By utilizing the internet as the primary medium, data is becoming more diverse and its volume is making it difficult for this conventional approach. A shift in paradigm is emerging in order to tackle this problem, and one of the emerging solutions is to utilize data streaming. In data streaming, data is generated continuously, often simultaneously, from a large number of data sources in small volumes, for example, some usual streaming candidates are an ecommerce purchase, social network information, sensor data, etc. In order to process these streams of data, specially designed systems are needed where each stream can be processed sequentially and incrementally to provide useful analytics.

One of the more interesting use-cases of data stream management includes the utilization of Internet of Things (hereinafter: IoT) technologies in smart cities. In a smart city environment, each user device or appliance can be regarded as a valuable data source for future stream analysis. These systems enable development of various smart city business models that leverage modern technologies, such as mobile and IoT, to improve the quality of life in new cities, and provide citizens ubiquitous and timely access to the right information and services.

In this article, we present a prototype of a system that enables collection of environmental data from mobile and IoT devices, storing and processing that data in scalable cloud infrastructure, and the delivery of the processed data to various clients.

Further text is organized as follows. Section 2 presents the theoretical background of the proposed system. Section 3 presents the architectural design of the system, while Sect. 4 presents the implementation details. Finally, we give concluding remarks.

2 Theoretical Background

The theoretical background section is organized into two parts:

- analysis of concepts and software architectures of IoT crowdsensing,
- analysis of data streaming aspects of IoT crowdsensing systems.

2.1 *IoT Crowdsensing*

As one method of crowdsourcing, crowdsensing refers to collecting data generated by sensors by a large group of people, i.e., the crowd [1–3]. It is typically done by

utilizing mobile device sensors, such as accelerometers, microphones, and others. Additionally, crowdsourcing can be done through various affordable IoT systems that can be owned and managed by individual citizens [3–5].

Recent works show [1, 6–9] that in order to be efficient, crowdsensing systems in smart cities need to collect data from various sources, combine and transform them, in order to provide relevant reports to all interested parties. For example, for the e-health stress management application, crowdsensing system should provide data on the noise, air quality, allergen levels, etc., and alert individual users if the stressors are detected. Such a system should be designed so as to be:

- Reliable—the system should be capable of serving its purpose timely,
- Scalable—the system should be capable to easily handle increased workloads,
- Distributed—various components are spread across multiple network locations.

In the past few years, microservice architecture has been recognized as an adequate approach for this type of system, having in mind the following characteristics [10, 11]:

- Microservices can be developed independently by small teams, allowing for faster and more efficient deployment of new features.
- Microservices can be implemented in diverse environments, allowing for ease of development.
- Resource sharing is discouraged, and data is designed with data streams in mind.
- The use of lightweight network protocols, such as HTTP, further simplifies development.
- The system is designed to work exclusively with events, and all communication is done by passing messages.
- The automation in service maintenance and use of devops methods and tools.

An example of microservice architecture for a crowdsensing system is shown in Fig. 1.

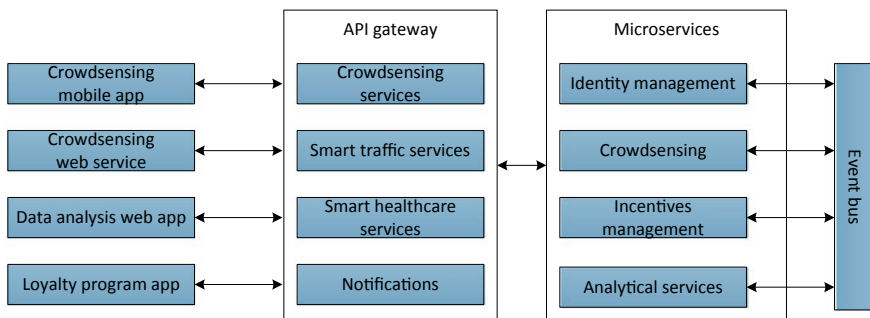


Fig. 1 An example of microservices in a crowdsensing system

2.2 Data Streaming in IoT Crowdsensing Systems

Crowdsensing systems generate data in data streams. Data streaming refers to continuously generated high volumes of data. Typical sources of data streams include sensors in IoT systems, event logs, clickstream, and similar. Data is usually semi-structured and ETL (extraction, transformation, load) is necessary before data analysis. For data stream management, specific software components capable of handling high volumes of data from various sources are necessary. This means the existence of adequate software infrastructure, not just a selection of one database or ETL tool. This infrastructure should provide:

1. capability for managing endless data streams,
2. near-real-time analysis, and
3. support for real-time pattern detection.

A core component of data streaming software infrastructure is a data streaming platform, Apache Kafka being the most commonly used.

Data streaming concept assumes that the events are recorded in a log (event stream), and clients can access this log and read any part of the data anytime.

In crowdsensing systems, data sources can be highly diverse, including IoT sensors, data from mobile devices, data from social media, etc. (Fig. 2). All of the received data is fed to the stream processor which translates messages from the senders' protocol to the receivers' protocol. The translated messages are then published within their respective topics. ETL tools and analytics software can be subscribed to these topic feeds and used to further process and analyze the data. The data collected by these streams can be stored in many different ways, such as:

- In a database or a data warehouse, such as MySQL, MongoDB, etc.
- In a message broker, such as Apache Kafka permanent storage.
- In a data lake in the cloud.

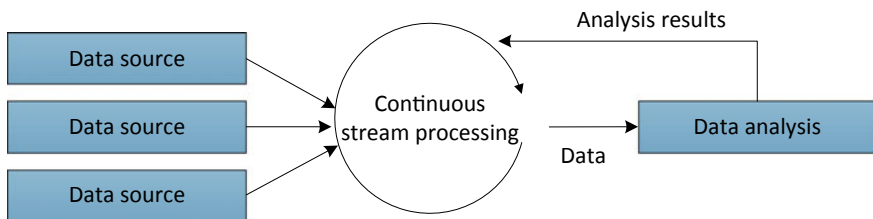


Fig. 2 Data streaming

3 Crowdsensing System Design

The main goal of the proposed system is to provide a comprehensive platform for crowdsensing in smart cities (Fig. 3). The system should be scalable in order to support participation of a large number of citizens in various crowdsensing initiatives. The system should support realization of different crowdsensing projects, such as collecting data about the environmental noise, vibrations, air quality, allergens, and others. The crowdsensing can be performed using mobile phones, or dedicated IoT stations with adequate sensors. Main features of the proposed system are presented in the Table 1.

The main components of the proposed system are:

- IoT devices are used for collecting data from sensors. If needed and possible, data measured from each sensor is processed and transformed at the edge device, and then sent to the cloud.
- Mobile applications for crowdsensing are used to collect measurements using participants' mobile devices. They provide collecting data such as noise at microlocations, vibrations in public transportation, number of people at a microlocation, etc.
- Data streaming infrastructure includes several components: (1) ingress service, i.e., a service that receives data from IoT systems and mobile applications; (2) stream processing cluster, which enables transformation of data streams; (3) access service which enables communication with consumers of the data, i.e., applications that use the data streams.
- Data analytics infrastructure provides services for real-time analytics. It enables implementation of advanced analyses, machine learning techniques, or other predictive models.

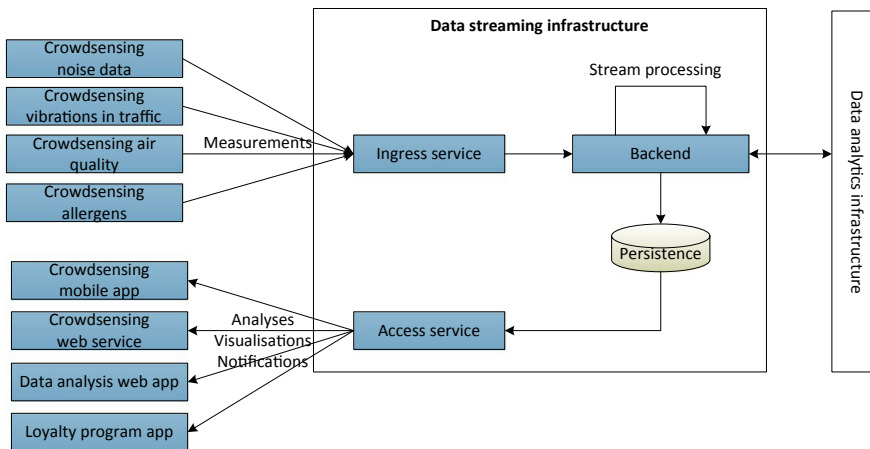


Fig. 3 The overview of the proposed system

Table 1 Main features of the proposed system

Feature	Description
User registration	Users who wish to participate in the crowdsensing can register to the system
Device or sensor registration	Registered users can connect their sensors, IoT systems, or mobile devices to the platform, in order to participate in the crowdsensing; one user can connect multiple data sources
Configuration	Users are able to configure the type of participation for each device (data stream). Participation can be opportunistic or participatory
Collecting measurements	Data measured by users' devices is streamlined into the crowdsensing system as an input stream
Measurement anonymization	Measurements have to be anonymized, before becoming available to other participants in the system. Still, the participation of each citizen will be monitored, in order to provide adequate incentives
Incentives management	The system allows tracking individual contributions and various incentives for participation
Data access	Access to individual measurements is provided to all authorized users
Data aggregation	Measured data is aggregated according to the aggregation criteria (type of measurement, location, time, etc.)
Notifications	The system allows configuration of notifications sent to data consumers
API access	Various client applications can access the data through the API
Administration and monitoring	All components of the system can be configured, managed, and are constantly monitored

- Consumer applications include various mobile, web, or other applications that consume crowdsensed data in any form. They usually present data using adequate visualisations and enable notifications.

4 Pilot Implementation

Within the pilot implementation, several components of the proposed system have been developed:

- IoT stations for gathering data about air quality, including the levels of certain allergens.
- Android mobile application for measuring traffic noise and vibrations in public transportation. More details are given in [6].
- Data streaming infrastructure based on Apache Kafka.
- A web application for visual presentation of measurement on a map.

- A mobile application for visual presentation of measurement on a map, as well as alerting users if high level of allergens was detected.

Figure 4 presents a component for crowdsensing the air quality data, using low-cost IoT stations that can be deployed with individual users. The sensors MQ2, MQ7, MQ135, or other are connected to Arduino board, and enable detection of various combustible gases, carbon-monoxide, ammonia, sulfide, benzene series steam, etc. Using MQTT protocol, the measure data is transferred to the RaspberryPi microcomputer, as the edge device that further sends data to the cloud using protocols such as REST or websocket. Some of the measurement parameters, such as frequency of measurements, can be configured remotely. Figure 4 illustrates the described IoT component of the system.

When the data is transferred to the data streaming infrastructure, it needs to be processed and transformed before being made available to consumers. An example of the data flow in the proposed system, implemented using Apache Kafka, is shown in Fig. 5.

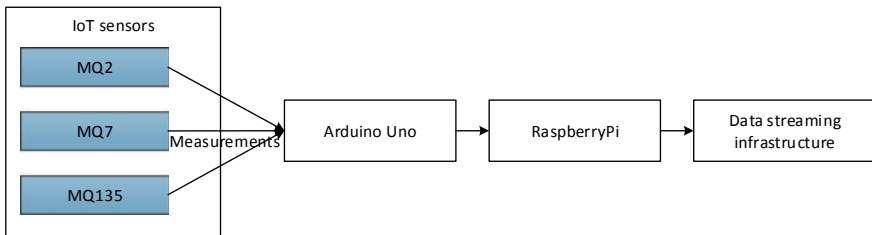


Fig. 4 Component for IoT crowdsensing air quality data

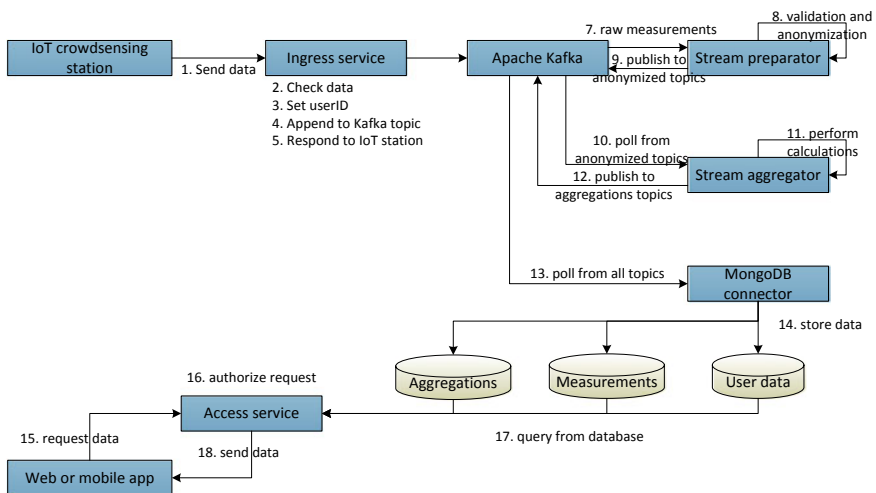


Fig. 5 An example of the data flow in the proposed crowdsensing system

Table 2 Selected list of topics used in the stream processing for CO₂ measurements

Topic	Description
raw_CO ₂ _measurements	Topic for entrance of air quality measurements
raw_CO ₂ _measurements_anonym	Anonymized air quality measurements
average_CO ₂ _time_location	Calculations of averages for a location at selected time periods; a topic is created for each combination of predefined time slots (daily, hourly) and selected locations of interest
max_CO ₂ _time_location	Maximal measured values for a selected time periods at a location

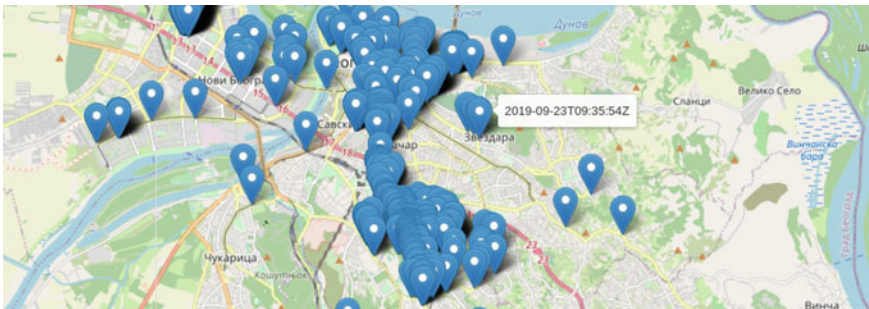


Fig. 6 All measurements for a selected parameter in the selected time period are shown in a web application

A selected list of Kafka topics from the crowdsensing system is shown in Table 2. Figure 6 presents an example of a report showing all measurements performed in the selected period.

Figure 7 presents a mobile application that provides alerts to its users about the allergens detected at the user’s location.

5 Conclusion

The primary goal of this paper is to offer an insight into a development of a stream-based IoT crowdsourcing platform. One of the primary requirements of the system was scalability and performance, which were achieved through Kafka stream processing alongside our own developed modules. The platform can be used to gather data on noise and air pollution, allergens, and other parameters in a smart city environment through various devices and sensors. The data is processed through Kafka and made available in real time through the web application and mobile app for the end users. The real-time nature of stream processing allows for interesting use cases in smart cities, where the designed system can offer user alerts in real time to people

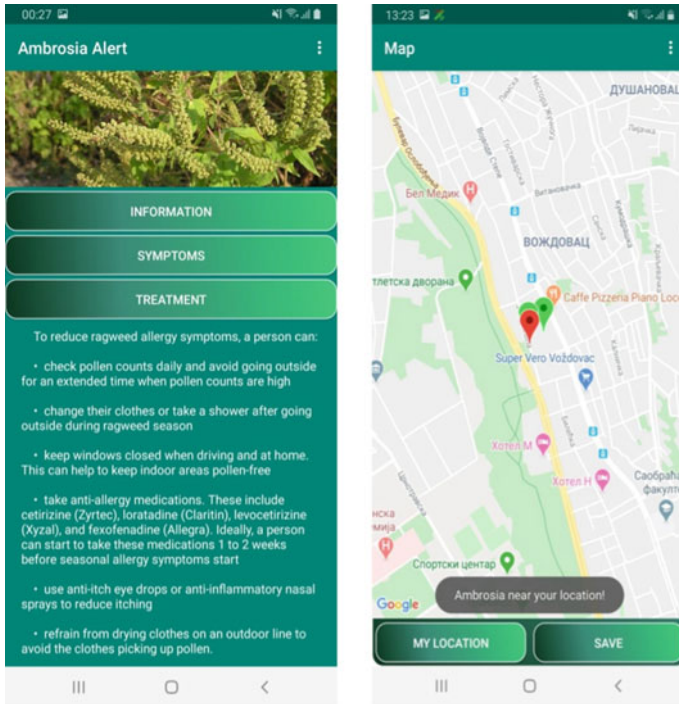


Fig. 7 Consumer mobile application for allergen alerting

affected by allergies or pollution. Some of the examples of the application of the system include, but are not limited to:

- Service for alerting citizens about the allergen level at their location
- Service for crowdsensing air quality data at microlocations, using IoT stations
- Service for crowdsensing noise data at microlocations, providing more detailed information than official systems
- Loyalty platform that provides citizens with various benefits for participating in data collection.

The model and the architecture presented can be utilized for most smart city services, but can easily be extended to accommodate other use cases. Additionally, the processed data the system provides can be utilized as a data source for advanced analytics by other systems such as Apache Spark. Further research will include a detailed analysis of performance of the proposed solution, a better optimization of Kafka topics with respect to the user queries, as well implementation of additional features within all the system components.

References

1. Labus, A., Radenković, M., Despotović-Zrakić, M., Bogdanović, Z., Barać, D.: Crowdsensing system for smart cities, In: Bećirović, D., Delić, H., (eds.) *Book of Proceedings 4th International Scientific Conference on Digital Economy—DIEC 2021*. Tuzla, Bosnia and Hercegovina (2021)
2. Lau, B.P.L., Wijerathne, N., Ng, B.K.K., Yuen, C.: Sensor fusion for public space utilization monitoring in a smart city. *IEEE Internet Things J.* **5**(2), 473–481 (2017)
3. Staletić, N., Labus, A., Bogdanović, Z., Despotović-Zrakić, M., & Radenković, B.: Citizens' readiness to crowdsourcing smart city services: a developing country perspective. *Cities* **107**, 102883 (2020)
4. Guo, B., Chen, C., Zhang, D., Yu, Z., Chin, A.: Mobile crowd sensing and computing: when participatory sensing meets participatory social media. *IEEE Commun. Mag.* **54**(2), 131–137 (2016)
5. Guo, B., Yu, Z., Zhou, X., Zhang, D.: From participatory sensing to mobile crowd sensing. In: *2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*, pp. 593–598. IEEE (2014)
6. Jezdović, I., Popović, S., Radenković, M., Labus, A., Bogdanović, Z.: A crowdsensing platform for real-time monitoring and analysis of noise pollution in smart cities. *Sustain. Comput. Inf. Syst.* 100588 (2021)
7. Kraft, R., Birk, F., Reichert, M., Deshpande, A., Schlee, W., Langguth, B., Baumeister, H., Probst, T., Spiliopoulou, M., Pryss, R.: Efficient processing of geospatial mhealth data using a scalable crowdsensing platform. *Sensors* **20**(12), 3456 (2020)
8. Stefanović, S., Milanović, S., Despotović-Zrakić, M.: Crowdsensing mobile healthcare application for detection of ambrosia. In: *XVII International Symposium Business and Artificial Intelligence, SYMORG Belgrade, September 7–9*, pp. 513–519. ISBN 978-86-7680-385-9 (2020)
9. Stefanović, S., Nešković, S., Rodić, B., Bjelica, A., Jovanić, B., Labus, A. : Development of a crowdsensing IoT system for tracking air quality, In: Despotović-Zrakić, M., Bogdanović, Z., Labus, A., Barać, D., Radenković, B. (eds.). *Book of Extended Abstracts of International Conference E-business Technologies EBT 2021*, pp. 117–119, Belgrade, Serbia. ISBN 978-86-7680-390-3 (2021)
10. Cerny, T., Donahoo, M.J., Trnka, M.: Contextual understanding of microservice architecture: current and future directions. *ACM SIGAPP Appl. Comput. Rev.* **17**(4), 29–45 (2018)
11. Chapman, M., Rasmussen, L.V., Pacheco, J.A., Curcin, V.: Phenoflow: a microservice architecture for portable workflow-based phenotype definitions. In: *AMIA Annual Symposium Proceedings*, vol. 2021, p. 142. American Medical Informatics Association (2021)

IoT Application Model in Secondary Education



Danijela Stojanović , Zorica Bogdanović ,
and Marijana Despotović-Zrakić 

Abstract The implementation of certain smart classroom services in the educational process makes it possible to increase the efficiency and quality of the teaching process, as well as the level of competence of students, with low financial investments and a short process of teacher training. The application of the Internet of Things and mobile technologies improves the quality of teaching, making it more interesting for students, thus technologies have significant potential for improving teaching and learning in all educational sectors. The advantage of this way of working is in the application of mobile technologies that students already use. Organizing the teaching process in smart learning environments enables quality and faster communication of students with teachers, active participation of students, and a better type of teaching as well as the efficient exchange of information in the teaching process. The research presented in the paper aims to point out the possibilities of improving the learning process in secondary education by applying pervasive technologies. The proposed model was evaluated with secondary school students.

Keywords e-learning · Secondary school · Internet of things · Pervasive technologies · Smart learning environments

1 Introduction

Internet of Things and mobile technologies are suitable for use for educational purposes because they do not require large infrastructure investments [1, 2]. Although an ideal smart classroom is expensive, a lot of functionality can be realized at a low cost [3–5]. Innovations can also be realized through new pedagogical approaches of teachers with small investments. Also, theoretical analyzes from the scientific literature and the results of previous research show that in higher education there

D. Stojanović (✉)
Institute of Economic Sciences, Belgrade, Serbia
e-mail: danijela.stojanovic@ien.bg.ac.rs

Z. Bogdanović · M. Despotović-Zrakić
University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia

are numerous experiences that can be adapted and applied to the improvement of secondary education [7, 8]. For developing countries, a better approach is the implementation of certain smart classroom services that would enable, with low financial investments and a short process of teacher training, to increase the efficiency and quality of the teaching process [9–11].

Analyzes of scientific research show that the Internet of Things (hereinafter: IoT) has great potential for application in formal secondary education [9, 12–14]. Their inclusion would improve the teaching process in secondary schools, and students would receive a modern and efficient education system, suitable for modern technological development.

Scientific research projects show that adequate equipping of secondary schools with modern technology is a prerequisite for effective improvement of teaching in secondary schools [15, 16]. In addition, teachers must be continuously educated, as well as develop and apply innovative e-learning methods [17].

Technologically advanced teaching aids enrich the educational process daily, creating smart learning environments. Smart learning environments are mainly based on IoT and mobile technologies [18]. IoT enables resource virtualization by presenting a concept that extends the Internet and the virtual world to physical things from the real world [19].

Smart environments are realized through various technologies, hardware, and software components, so the IoT value chain itself consists of several different categories: sensors, actuators, radio/communication chips, microcontrollers, modules, software platforms, application software, telecom infrastructure, service infrastructure [20].

The goal of this paper is to point out the possibilities of improving the learning process in secondary education by using pervasive technologies in the learning process.

2 Literature Review

2.1 *Electronic Education in Secondary Schools*

Newer generations of secondary school students view the world through the prism of computers, mobile phones, cameras, players, video games, and other modern information technology devices, which give them the ability to perform multiple tasks at once and at the same time listen to music, surf the Internet or chat online while doing homework [21].

With the technological development of society, the importance of education is growing, which thus becomes a condition for the survival and progress of modern society [11, 15]. In line with that, there is an increasing need for educational activities to be realized regardless of location and time. However, the question is whether the standard educational process can meet such requirements. The application of modern technologies in the educational process can certainly enable that [16].

The creation of smart learning environments has also improved e-learning, which is being upgraded by using technologies of smart learning environments. Smart education is considered to be education supported by technologies, smart tools, and devices. Specific concepts of e-learning are collaborative and mobile learning, which have the potential to support the development and application of crowdsourcing in smart learning environments [17].

2.2 Technologies of Smart Learning Environments

A smart environment is an ecosystem composed of objects, i.e., sensors, and actuators that interact with each other and can manipulate and process large amounts of data to automate actions that are often repeated [18, 19]. IoT enables the virtualization of resources by presenting a concept that extends the Internet and the virtual world to physical things from the real world. Physical things have the possibility of virtual representation by connecting to the Internet and unambiguous identification. IoT can also be defined as a higher level of pervasive technologies and intelligence, where components, products, services, and platforms are connected, and everything is integrated into a communication network for digital processing. IoT is based on several different disciplines and technologies, including sensors, communication technologies, semantic and security technologies, but also requires a specific configuration for object identification, lightweight protocols, open/closed data sharing, etc. [20].

The management of a smart learning environment is based on core values, such as human rights, ethics, the common good, legislation, and transparency of educational procedures [21].

Smart environments are realized through various technologies, hardware, and software components, so the IoT value chain itself consists of several different categories [22]:

1. Sensors—generate most data;
2. Actuators—perform actions;
3. Radio/communication chips—enable connectivity;
4. Microcontrollers—process data;
5. Modules—combine radio, sensors, and microcontrollers and allow them to be added to the device;
6. Software platform—enables management and capabilities of the IoT network;
7. Application software—provides the collected information in a format that is suitable for end-users, so that they can use and analyze it;
8. Telecom infrastructure—enables mainly wireless data transmission;
9. Service infrastructure—enables design, installation, monitoring, and servicing of IoT implementation.

The most famous smart environments are smart classrooms. In smart learning environments, several different technologies are used to develop crowdsourcing models. The model is being developed in smart learning environments, so it is necessary to

use IoT technologies. The development of crowdsourcing systems requires the use of web and internet technologies, as well as technologies for building portals [9, 23].

Portals bring together information from different sources and present it to portal visitors. In this way, different goals can be achieved, which relate to informing users, advertising, connecting users, etc. Today’s portals are even more important because they enable user interaction and their direct involvement [15, 24].

3 Model Development

Modeling the environment for secondary education is a way to connect the services of the IoT into a unique and planned organized technological, organizational, and educational system, aimed at the overall improvement of secondary education. The proposed model of e-education in smart learning environments includes the following components: e-education system framework, smart classroom architecture, smart classroom software infrastructure, integration of IoT infrastructure components with e-education system, e-education system services, and new pedagogical approaches in e-education (Fig. 1).

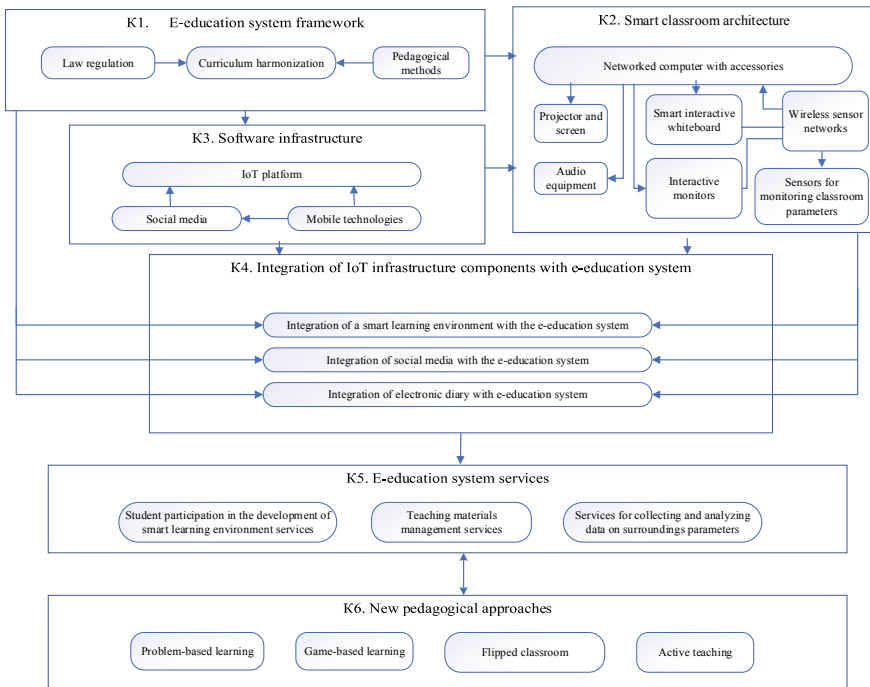


Fig. 1 General structure of the model of secondary school e-education based on the IoT

The e-education system framework is the first component in the structure of the secondary school e-learning model based on IoT technologies. It includes law regulations, curriculum harmonization, and pedagogical methods. Implies harmonization of law regulations and pedagogical methods with curricula.

The next component is **smart classroom architecture**, which involves designing a smart learning environment for secondary school education.

The smart classroom represents a projected smart learning environment for secondary education [9].

When creating smart classrooms, it is necessary to make a compromise between the initiatives and wishes of teachers, students, and parents. The smart classroom should provide easy access to information via an intranet, internet, electronic diary, assessment and learning management system, digital content, school website, then through individual or specific content made for individual departments or classes, as well as through blogs, wikis, and social media that are used in school and are easily accessible.

To design a smart learning environment for secondary education, it is necessary to have in the classroom: networked computer with accompanying equipment, smart interactive whiteboard, interactive monitors, projector and screen, wireless sensor networks, audio equipment, sensors for monitoring classroom parameters.

By measuring parameters related to the physical environment, such as lighting, air quality, noise, or room temperature, student satisfaction can be accurately determined during class. Using various devices or mobile phones with sensors, some of the parameters of the physical environment can be easily measured. Based on these parameters, a smart classroom can be designed, which will be able to analyze the parameters of the physical environment to determine student satisfaction with the quality of teaching at a given time. Thus designed and implemented smart classroom systems, according to research, can determine how satisfied students are with the quality of lectures with an accuracy of more than 93% [25].

The software infrastructure of the smart classroom consists of the IoT platform, social media, and mobile technologies.

IoT platforms should integrate heterogeneous devices, as well as store and manage data collected from sensors. To achieve a high degree of scalability, reliability, redundancy, and better use of available computing resources, it is desirable to use a cloud platform. Most elements of the IoT platform are cloud-based and wirelessly connected by a combination of different technologies, including mobile technologies, web services, and the like [26, 27].

When it comes to IoT devices that can be used in secondary school education, Raspberry Pi microcomputers (RPI) and Arduino microcontrollers appear to be good solutions. Using these devices, IoT solutions that do not require specialized components or expert knowledge for implementation can be easily created [28].

Some examples of open source IoT platforms that can be used in education are the *ThingSpeak Platform* and the *Mainflux Platform*.

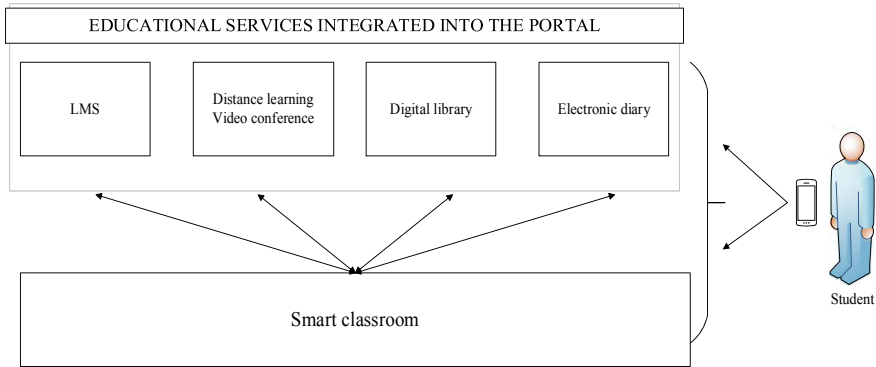


Fig. 2 Educational services in a smart learning environment

Both platforms are suitable for use in a smart educational environment. From the aspect of economic profitability, it is important that the implementation, implementation, and development of a smart environment are simple and do not require large costs.

Within the software infrastructure, there are social media and mobile technologies that are connected to the IoT platform. The paper [24] presents an approach to the development of a web portal for adaptive e-learning. Mobile and wireless technologies have the task of enabling communication and cooperation between intelligent devices, mobile devices, and the Internet [29].

The integration of IoT infrastructure components with the e-education system is the fourth component in the model. It represents the unity of the work of the previous three explained components as a whole. It consists of several parts: integration of a smart learning environment with the e-education system, integration of social media with the e-education system, integration of the electronic diary with the e-education system (Fig. 2).

Today, learning is no longer limited to the school and the classroom, and educational institutions provide a large number of extracurricular activities to improve students' knowledge and skills. But, while earlier information was obtained in libraries or computer rooms, now it can and must be available to everyone regardless of place and time.

By setting up *WiFi* access points in the school, students are given access to the school computer network and the Internet. Within the school, areas should be created where students will have the opportunity to use digital tools continuously, and materials for classes should be published on the school portal, which students could access and download.

E-learning system services include student participation in the development of smart environment services, services for managing teaching materials, services for collecting and analyzing data on environmental parameters.

Students can develop their services and share them with other students, and by integrating student services with cloud-based services and external web APIs, more complex services can be assembled.

Creating reports based on reading data from sensors is a typical scenario in applications of Internet intelligent devices, so students are provided with a simple API for this purpose. It allows the storage, use, and sharing of data read from sensors, as well as the input or generation of external data, to stimulate some specific conditions, which cannot be directly measured in the educational environment.

For the e-learning model based on the IoT to be successfully established in secondary education, teaching equipment and infrastructure are needed. The limiting factor in this may be financial resources, given that it is necessary to provide equipment for the design and implementation of various services of IoT.

Equipping classrooms with smart teaching aids is organized on two levels. The first level is global and includes the implementation of a solution that will be used by all classroom users. The second level is individual and involves equipping workstations [26].

The last component in the model is **new pedagogical approaches in e-education**.

Examples of some of the pedagogical models that can be applied in secondary school teaching are the following: Problem-based learning, Game-based learning, Flipped classroom, Active teaching [7, 30–32].

4 Evaluation

The evaluation of the presented model was conducted with the students of the Secondary School of Economics in Belgrade. The results of the evaluation are presented in the article [9], and they show that the implemented model of e-education based on IoT devices has enabled an increase in the level of knowledge and interest of students to learn using smart devices and mobile phones in the process of knowledge assessment.

Following the example of the proposed model, the research was conducted in higher education, with students of the Faculty of Organizational Sciences, University of Belgrade. A smart learning environment system has been developed for the research. The evaluation is presented in an article [23], and the results obtained showed that students achieve better results by using modern technologies in the learning process compared to traditional ways of learning. Increased motivation, as well as a high level of interest, is relevant to the way of working with the application of pervasive technologies.

The conducted research indicates that new methodological approaches to teaching can contribute to better vertical integration of secondary and higher education in the educational process.

5 Conclusion

The main goal of the research is the efficient and effective improvement of teaching in secondary schools by applying the concepts and technologies of the IoT. As part of the evaluation of the model, testing and measurement of relevant parameters were realized.

One of the preconditions for the efficient use of pervasive technologies in education is the adequately equipping secondary schools with modern equipment and technology. After that, continuous theoretical and practical education of teachers is necessary, so that they first understand and accept the role of pervasive technologies in the teaching process, and then they can pass it on to students.

It is quite certain that in that way, the educational process in secondary schools would be significantly improved, enhanced, and raised to a much higher level, and students would get a modern and more efficient education system suitable for modern technological development.

Acknowledgements The authors are thankful to the Ministry of education, science and technological development, project numbers 11143.

References

1. Aduwa-Ogiegbaen, S.E., Iyamu, E.O.S.: Using information and communication technology in secondary schools in Nigeria: problems and prospects. *J. Educ. Technol. Soc.* **8**(1), 104–112 (2005)
2. Cakir, R., Yildirim, S.: Who are they really? A review of the characteristics of pre-service ICT teachers in Turkey. *Asia-Pac. Educ. Res.* **24**(1), 67–80 (2015)
3. Asterhan, C.S., Rosenberg, H.: The promise, reality and dilemmas of secondary school teacher–student interactions in Facebook: the teacher perspective. *Comput. Educ.* **85**, 134–148 (2015)
4. Salleh, S.M., Laxman, K.: Investigating the factors influencing teachers' use of ICT in teaching in Bruneian secondary schools. *Educ. Inf. Technol.* **19**(4), 747–762 (2014)
5. Sanders, M., George, A.: Viewing the changing world of educational technology from a different perspective: present realities, past lessons, and future possibilities. *Educ. Inf. Technol.* **22**(6), 2915–2933 (2017)
6. Beetham, H., Sharpe, R.: An approach to learning activity design. In: *Rethinking Pedagogy for a Digital Age: Designing for 21st Century Learning*, pp. 26–40. Routledge, Taylor & Francis Group (2013)
7. Stojanović, D., Bogdanović, Z., Despotović-Zrakić, M., Naumović, T., Radenković, M.: An approach to using Instagram in secondary education. In: *The 14th International Conference on Virtual Learning ICVL*, pp. 247–253 (2019)
8. Michailidis, N., Kapravelos, E., Tsiatsos, T.: Examining the effect of interaction analysis on supporting students' motivation and learning strategies in online blog-based secondary education programming courses. *Interact. Learn. Environ.*, 1–12 (2019)
9. Stojanović, D., Bogdanović, Z., Petrović, L., Mitrović, S., Labus, A.: Empowering learning process in secondary education using pervasive technologies. *Interact. Learn. Environ.* **1**(1), 1–14 (2020)
10. Stojanović, D., Domazet, I.: Use of information technologies in educational purposes-case from Serbia. *Econ. Anal.* **53**(2), 68–78 (2020)

11. Stankić, R., Jovanović-Gavrilović, B., Soldić-Aleksić, J.: Information and communication technologies in education as a stimulus to economic development. *Ekon. Horizonti* **20**(1), 61–73 (2018)
12. Glaroudis, D., Iossifides, A., Spyropoulou, N., Zaharakis, I.D., Kameas, A.D.: STEM learning and career orientation via IoT hands-on activities in secondary education. In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 480–485 (2019)
13. Mavroudi, A., Divitini, M., Gianni, F., Mora, S., Kvittem, D.R.: Designing IoT applications in lower secondary schools. In: *IEEE Global Engineering Education Conference (EDUCON)*, pp. 1120–1126 (2018)
14. Palagolla, W.W.N.C.K., Wickramarachchi, A.P.R.: Effective integration of ICT to facilitate the secondary education in Sri Lanka. *arXiv Prepr. arXiv1901.00181* (2019)
15. Stojanović, D.: Analiza realizacije učenja na daljinu u Srbiji za vreme pandemije virusa COVID 19. In: *Black swan in the world economy 2020*. Institut ekonomskih nauka, Beograd, pp. 121–140. ISBN 978-86-89465-54-9 (2020)
16. Maletić, M., Barać, D., Naumović, T., Bogdanović, Z., Radenković, B.: Blending Crowdvoting in Modern e-Learning Environments. *Int. Rev. Res. Open Distrib. Learn.* **20**(2) (2019)
17. Zahirović Suhonjić, A., Despotović-Zrakić, M., Labus, A., Bogdanović, Z., Barać, D.: Fostering students' participation in creating educational content through crowdsourcing. *Interact. Learn. Environ.* **27**(1), 72–85 (2019)
18. Palma, D., Agudo, J.E., Sánchez, H., Macías, M.M.: An internet of things example: classrooms access control over near field communication. *Sensors* **14**(4), 6998–7012 (2014)
19. Petrović, L., Jezdović, I., Stojanović, D., Bogdanović, Z., Despotović-Zrakić, M.: Development of an educational game based on IoT. *Int. J. Electr. Eng. Comput.* **1**(1), 36–45 (2017)
20. Lakshmi, I.: A vision, architectural elements, and future direction of internet of things (IoT). *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **3**(1), 306–328 (2018)
21. Garcia, A., Marquez, J., Solarte, Z., Villanueva, J.: IoT in education: integration of objects with virtual academic communities. In: *New Advances in Information Systems and Technologies*, vol. 444, pp. 201–212. Springer, Verlag (2016)
22. Friess, P., Riemenschneider, R.: New horizons for the Internet of Thing in Europe. In: *Building the Hyperconnected Society—IoT Research and Innovation Value Chains, Ecosystems and Markets*. River Publishers, Denmark (2015)
23. Petrović, L., Stojanović, D., Mitrović, S., Barać, D., Bogdanović, Z.: Designing an extended smart classroom: an approach to game-based learning for IoT. *Comput. Appl. Eng. Educ.* (2021)
24. Radenković, B., Despotović-Zrakić, M., Barać, D., Bogdanović, Z., Milić, A.: Web portal for adaptive e-learning. In: *2011 10th International Conference on Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS)*, pp. 365–368 (2011)
25. Gligoric, N., Uzelac, A., Krco, S., Kovacevic, I., Nikodijevic, A.: Smart classroom system for detecting level of interest a lecture creates in a classroom. *J. Amb. Intell. Smart Environ.* **7**(2), 271–284 (2015)
26. Bogdanović, Z., Simić, K., Milutinović, M., Radenković, B., Despotović-Zrakić, M.: A Platform for Learning Internet of Things. *Int. Assoc. Dev. Inf. Soc.*, 259–266 (2014)
27. Perry, M.J. *Evaluating and Choosing an IoT Platform*. O'Reilly Media (2016)
28. Milutinović, M., Simić, K., Labus, A., Bogdanović, Z., Despotović-Zrakić, M.: Platforma za učenje interneta inteligentnih uređaja. In: *XIII međunarodni naučno-stručni simpozijum INFOTEH-JAHORINA 2014*, vol. 13, pp. 759–762 (2014)
29. Basole, R.C., Karla, J.: On the evolution of mobile platform ecosystem structure and strategy. *Bus. Inf. Syst. Eng.* **3**(5), 313–322 (2011)
30. Petrović, L., Stojanović, D., Labus, A.: Development of an educational game: Augmented reality approach to edutainment. In: *XVI International Symposium Doing Business in the Digital Age: Challenges, Approaches and Solutions SymOrg 2018*, pp. 96–107 (2018)

31. Maenpaa, H., Varjonen, S., Hellas, A., Tarkoma, S., Mannisto, T.: Assessing IOT projects in university education—a framework for problem-based learning. In: Proceedings—2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering and Education Track, ICSE-SEET 2017, pp. 37–46 (2017)
32. Ivanović, B.: ‘Flipped classroom’: pedagogical model that changes the concept of traditional teaching. *Ekon. signali Posl. Mag.* **13**(1), 57–66 (2018)

Digital Marketing Strategy: A Step-By-Step Approach



M. J. Guerra  and F. O. Silva 

Abstract Digital communication has had a strong impact on the development of society. Investing in multimedia to produce original content as a means of promotion of a company has become crucial. The creation of multimedia content for social media platforms, the existence of a website, and the dissemination of information in digital format is an essential criterion to inform the mission, goals, products and services that a company intends to provide to its current or potential customers. The definition of a digital marketing strategy is central to give visibility to the company and strengthen the relationship with its target audience. The main purpose of this paper is to present the digital marketing strategy of “Alpha” (assumed name)—a Portuguese technology company specialized in Internet-related services and products. For this purpose, data were collected from a qualitative survey applied to companies, in the same sector of activity. The aim of this study is to understand and to identify the methods, procedures, and tools that companies are using to promote their services. A semi-structured interview was conducted with the CEO of the company intending to understand the company’s perception about the key factors to elaborate a digital marketing strategy. Findings based on the questionnaire and on the depth interview indicated that most of the participant companies have developed a digital marketing strategy. This study reinforces the idea that complementing market analysis based on digital marketing with the production of content through multimedia technologies provide benefits for the company’s success.

Keywords Digital marketing strategy · Transmedia communication · Multimedia content

M. J. Guerra (✉)
ESMAD—P.PORTO, Vila do Conde, Portugal

F. O. Silva
ISLA—IPGT, Research and Development Centre, Gaia, Portugal
e-mail: firmino.silva@islaguia.pt

1 Introduction

New technologies, in particular the growth of the internet, have changed the way individuals and companies communicate and interact, making communication faster and more efficient. More specifically, the use of new online tools has made it possible to promote the image of the company and its services or products, creating targeted advertising campaigns and attracting new customers [3, 4].

In this context, to strengthen the communication between companies and their public, it is essential to create a digital marketing strategy that recognizes the importance of interactivity with the public to gain new customers and ensure the maintenance of the current ones.

The selection of “Alpha”, as the entity to be studied in this study, gathered some relevant features, such as: being a growing start-up, found by a young team, dedicated to new technologies, with a vision and mission focused on promoting the digital literacy, making companies more digital and inclusive.

This paper aims to contribute to the definition of Alpha’s digital marketing strategy and the consequent implementation of multimedia content for promotion on social media platforms. The starting question of this study is how to enhance the promotion of Alpha’s technological products, through a digital marketing strategy based on specific multimedia objects?

The methodology applied used a qualitative and quantitative approach. To carry out the study focused on the present, regarding the definition of the digital marketing strategy for “Alpha”, two phases of data collection were created. First, we approached the universe of participants through the application of a questionnaire submitted online, from which we obtained a sample of 30 companies. In a second phase, we conducted a semi-structured interview with Alpha’s CEO.

This paper is structured as follows: the state of the art on Digital Marketing and multimedia technologies, which allowed upstream to study and define a digital marketing strategy for “Alpha”. Methodological options that guided the study; results obtained by data collection; finally, a reflection about main findings and future work proposals.

2 Digital Marketing, Social Media and Multimedia Technologies

We can consider that marketing is a planning process that seeks the total satisfaction of consumers, through a clear forecast of what they need and want, in a timely, and need [7]. Along the same lines, Armstrong and Kotler [1] consider that the main objective of marketing is to attract customers, providing them with the satisfaction of their needs. Patrutiu-Baltes [15] argues that in the case of traditional marketing, the marketing strategies of companies are divided between the consumer and the product, on the other hand, in digital marketing, the focus is on the consumer and

the identification of their needs to achieve effective communication between the company and the consumer.

Reflecting on the various definitions, which do not differ from each other, we can conclude that marketing is the set of methods and means a company must create, communicate, identify products or services that meet tastes, motivations, and preferences of current and potential customers, satisfying their needs and leading to the acquisition or consumption of these products/services. It is shown as a dynamic area that is attentive to the requirements and needs of companies and consumers, a relationship that can be enhanced if we add the added value of technological innovation and the virtual world.

2.1 Digital Communication

The consumption of the digital age has brought new perspectives and new ways of using marketing to identify and position a brand or company, communicate with the market, and attract customers [21].

The role of digital communication is widely spread as it is possible to adapt its content to different social platforms and its audience, making communication more personalized and interactive. According to Martínez-Rodrigo and Sánchez-Martín [13], the development of Information and Communication Technologies (ICT) and the adoption and implementation of the Internet have caused substantial changes at the technological and social level, which, in turn, have favored the emergence of new communication platforms. These new environments, created thanks to Web 2.0, culminated in a global transformation that led to substantial changes in the way companies and brands communicate with their audiences.

In this regard, for Jenkins [9], transmedia narratives are important in the ability of a story to expand through different media. In turn, for Scolari [20] they are stories told in different media. The most significant stories tend to flow across multiple media and communication platforms. A transmedia narrative structure can be considered as such when different languages (verbal, audiovisual, iconic, etc.) and different media (cinema, television, radio, etc.) are present, that is, it is not an adaptation of one medium to another, but the narrative can be seen as a whole, told from different media and languages.

Communication is a fundamental part of the recognition and prestige of a brand. The communication and interaction process between a company and their public has become an asset in an increasingly digital society.

2.2 The Role of Online Social Networks in a Company's Communication

With the evolution of ICT, online social networks had very strong growth. Most companies are present in the online world and have a “profile” on the appropriate digital platforms for their business area. The social networks facilitate not only the interpersonal relationship between users, but also the interaction of these subjects with brands and social media. Online social networks influence all areas of society and are used on a personal, professional, or academic level [13].

According to the report presented by Hootsuite [8], on “Social Trends 2020”, the world population is currently 7.75 billion, of which 5.19B have mobile phones, 4.54B are Internet users and 3.80B are active users of online social networks. Regarding data on Portugal, whose total population is around 10.21 million, 8.52M use the Internet and 7M are present on online social networks. Through these numbers, it is easy to see that despite having a strong presence in the digital world, they also demonstrate great growth potential and an opportunity for companies to create and strengthen their visibility in an increasingly digital market, where personalized communication is privileged, targeted and interactive, between a company and its target audience.

2.3 Implementing a Digital Marketing Strategy Based on Multimedia Technologies

Justifying the importance of a company implementing a digital marketing strategy, Ryan [17] states that formulating such a strategy will help to make informed decisions about its foray into the virtual market and ensure that efforts are focused on the most relevant elements for the business area. It is crucial to understand how the digital market works, how it evolves, and how it relates to the company and current or potential customers. Through digital channels it is possible to connect to a much wider audience, transcending issues such as geography or time zones, as well as adjusting communication to a particular market niche.

Technological advances have strongly influenced marketing, such as the Internet, social networks, AI, IoT, AR, VR, or Video Mapping. Deepening these aspects, it is expected that there will be changes in consumer preferences and behavior, not only because of the way in which production and distribution are carried out, but also in relation to the way in which communication and interaction between individuals, is carried out—the use of mobile technology has increased this capacity.

The successful implementation of the digital marketing strategy, foresee appealing multimedia content for the public, meeting the audience, to create a relationship, and mesmerize them.

The creation of innovative and impactful content for the public can be achieved using different technologies, such as motion graphics, augmented reality, or video mapping.

Motion Graphics

Motion graphics include moving, rotating, or resizing images, videos, and text over time on the screen, usually accompanied by a soundtrack (voice, music, etc.). They tend to resort to simplification and abstraction, reducing an image to a diagrammatic form, not meaning that they cannot use other graphic elements. Any visual element (text, image, texture, shape, or line) is suitable to be part of a motion graphic. In this sense, the characteristics of a motion graphic are based on time and movement, it aims at communication, rather than simply being a visual experience, it is like an extra layer of information that helps to explain a point of view or a concept, be it simple or complex [5].

Augmented Reality

According to Schmalstieg and Höllerer [18], augmented reality holds the promise of creating automatic and actionable links between the physical world and electronic information. Provides the user with a simple and immediate interface to an electronically enhanced physical world. On the other hand, according to Aukstakalnis [2] the term “augmented reality” is generally applied to a variety of display technologies capable of super-imposing or combining alphanumeric, symbolic, or graphical information with the user’s view of the real world. That is, these alphanumeric or graphical enhancements would be aligned, correlated, and stabilized within the user’s real-world view in a spatially contextual and intelligent way.

Video Mapping

Video mapping is the technique that consists of video projecting in small or large objects, building facades, or even on the human being’s body. With specific tools, it is possible to build, virtually, objects of two or three dimensions, representing in the best possible way the physical object on which we see the projection being carried out. In addition to this aspect, in these computer programs we create the narrative to be projected, combining different types of media, such as video, audio, images, text, animation [6, 12].

Combining digital marketing with the use of this type of technology is an excellent way to capture the public’s attention, provide interaction between the company and the consumer, and promote, publicize, and disseminate a product or service in an original way.

3 Methodology

Systematizing hypotheses, it is possible to anticipate five scenarios:

1. The multimedia objects that best fit in the promotion of Alpha's technological projects are: (a) Videos and/or still images; (b) Content in motion graphics.
2. The technologies that best fit the promotion of Alpha's technological projects on social media platforms are: (c) Videos and illustration; (d) Motion graphics.
3. The most innovative multimedia content adapted to each social network platform will attract more followers and the consequent brand awareness.

The first step was to build a database that aimed to identify and organize information about companies by sector of activity. We focused on entities that provided the consumer with services such as development of applications or services for the Web and/or mobile; website creation—services like “Alpha”. We checked the website of each company presented, identifying the services provided. Through the information we found available, it was possible to select the companies that provide the services we wanted, resulting in a database with company name, telephone number and e-mail, the URL, and their geographic location.

3.1 Questionnaire

The construction of the questionnaire was designed to be essentially an instrument for characterizing the methods, procedures, and tools used by the companies to promote services.

The questionnaire consisted of close-ended questions except for the “other” option and observations. The questionnaire completion time was relatively short (maximum 5 min) and was submitted online. The target sample group was contacted by email and was invited to respond to the questionnaire online.

3.2 Interview

Data was also collected through a semi-structured interview. The interview guide included close and open-ended questions. Questions focused on Alpha's perception about what they intend to achieve with the digital marketing strategy; the promotion of products currently carried out and how they intend to do it in the future; what they want to achieve through the website and portfolio.

3.3 Sample

The first phase of the study identified all companies in the sector of activity: other activities related to information and information technologies. According to data from the National Institute of Statistics (INE), in 2019 there were 2018 companies in Portugal (Instituto Nacional de Estatística 2021).

To build a sample according to the outlined objectives, exclusion criteria were defined to homogenize the services provided by the companies.

Thus, all companies that did not provide services related to the development of applications/services for the Web and all those connected to the development of applications/services for mobile devices and/or the creation of websites were excluded.

For phase one of the study, the application of the questionnaire was aimed at companies that met the criteria defined above, totaling 126 companies in Portugal: 43 in the North region, 26 in the Center, 49 in the Metropolitan Area of Lisbon, 4 in Alentejo, 4 from Algarve and none from the Autonomous Regions of the Azores and Madeira.

For the interview study, a key informant was contacted and invited to be interviewed. The Alfa's CEO was knowledgeable or well-briefed on the interview themes.

3.4 Procedures and Information Processing

This study is organized into three phases with different requirements in terms of procedures. The first phase intended to identify, through information available on the Internet, companies belonging to the same sector of activity as "Alpha" and whose services available to customers are development of applications/services for the Web, development of applications/services for mobile devices and/or creating websites.

After preparing the questionnaire, the second phase of the study began with the sending of 126 emails to the selected companies. The email dealt with the request to participate in this phase of the study, explaining the purpose of the study, the filling instructions, and the commitment to secrecy and confidentiality regarding the information provided.

Data were collected between February 23 and April 28, 2021. The final sample included 30 participants (23.81%).

Data were processed and analyzed with proper software. The survey was analyzed in the SPSS statistical program, version 26. After conducting the interview, we transcribed it and wrote an intelligible speech including punctuation and suppressing the useless elements of language.

Participating companies were not identified. The data collected through the interview could be reviewed by the participants before being used in the final version. The secrecy and confidentiality of data are guaranteed.

4 Results

4.1 Questionnaire

Findings obtained through the questionnaire applied to the companies (Table 1).

4.2 Interview Study

Analyzing the interview, we verified that all products, services, or solutions developed by “Alpha” are people-centered, always meeting clients’ needs. The implementation of a digital marketing strategy is essential to establish a set of guidelines that facilitate the dissemination and promotion of the “Alpha” brand and its products or services. We also found that it is necessary to improve the promotion of Alpha’s brand, products, and services gradually and continuously. The use of social networks and content adapted to each platform is also considered essential.

Alpha’s digital marketing strategy is based on 11 steps (Fig. 1).

Briefly describing Fig. 1, we started by presenting a description of the company, followed by the objectives and target audience, creating three personas aligned with it. We defined a budget, analyzed the internal environment using SWOT analysis and the competitive environment through the analysis of questionnaire.

In step two, about the website, portfolio, and Blog, we suggest tools to obtain reports on website performance, as well as points for improvement. For example, in the case of the Blog, the published contents would gain if they were shared on all social media platforms used by the company. It would be a way for potential customers to have knowledge and a better perception of the contents published on the Blog.

In step three, on multimedia contents, we identified the importance of adapting multimedia contents to the characteristics of each social media platform. Thus, favoring transmedia communication, through the publication of content in various media in a complementary way.

Step four refers to social media platforms, where we identified the best times to publish content. It is recommended to publish content at times with the greatest impact on the target audience of each social network.

We defined step five, on Google Marketing, identifying the most suitable Google tools for creating advertising campaigns, analyzing metrics, and optimizing the website. We also presented the main aspects that impact Google’s positioning.

Table 1 Questionnaire results applied to the companies

<i>The company</i>	
Services provided by companies to customers (most prevalent)	<ul style="list-style-type: none"> – Development of applications or services for the web (23.4%) – Website creation (21.8%) – Graphic design (16.1%) – Development of applications or services for mobile devices (15.3%) – Digital marketing (14.5%)
Means of promotion (most used)	<ul style="list-style-type: none"> – Website (32.3%) – Social networks (26.9%) – Email marketing (15.1%) – Newsletter (9.7%)
<i>The use of digital marketing strategies</i>	
Contexts they use digital marketing strategies	<ul style="list-style-type: none"> – Digital marketing strategies for the company (52.5%) – Digital marketing strategies for each product or service (35.0%) – Do not use any type of digital marketing strategy (12.5%)
<i>Social media platforms used</i>	
The social media platforms (most used)	<ul style="list-style-type: none"> – Facebook (73.4%) – Instagram (70.0%) – LinkedIn (56.7%)
Content publications are similar across all social media platforms	<ul style="list-style-type: none"> – Published content is not the same (66.7%) – It is similar (33.3%)
<i>The interaction between the company and its followers</i>	
Interaction with their followers	<ul style="list-style-type: none"> – Interact (83.3%) – Do not perform any type of interaction (16.7%)
<i>The typology of multimedia contents published on social media platforms</i>	
Typology of multimedia content published (most used)	<ul style="list-style-type: none"> – Images (30.2%) – Text (26.0%) – Videos (21.9%)
Technologies for creating content to be published (most used)	<ul style="list-style-type: none"> – Illustration/images (39.4%) – Videos (31.0%) – Motion graphics (14.1%)

For step six, on mobile devices, we presented a set of websites that provide free diagnostics on website performance on mobile devices.

In step seven, regarding opportunity marketing, we identified the importance of it and of reacting immediately. We identified some Portuguese companies that practice it exceptionally, such as “Licor Beirão”, “Super Bock” and “Dr Bayard”.

Moving on to step eight on advertising, we identified the most appropriate advertisements for the company and on which social media platforms.



Fig. 1 Alpha’s digital marketing strategy

Step nine, on Search Engine Marketing (SEM), refers to the importance of resorting to SEM when investing in advertising, since the results will be less efficient if it is not optimized.

Step ten presents the four main digital marketing metrics used: ad performance metrics; site traffic capture metrics; e-commerce sales funnel metrics and digital marketing platform metrics.

In the last stage, we identified other information. This step can be adapted depending on the multimedia content, the ads that will be invested, and the marketing opportunity that arises.

5 Conclusion

This study reinforces the relevance of developing digital marketing strategies. We live in a digital age and in all domains access to computers, Internet and social networks has reconfigured the way of life of individuals and companies.

As noted in this study, the use of new technologies and the consequent presence of companies and consumers on different social media platforms has strengthened communication and the relationship between a company and their customers. This communication is characterized by being personalized, interactive, and directed to a specific target audience. Regarding multimedia content, there is an increasing focus on the creation of innovative products, which attract the public’s attention and are adapted to each social network platform.

For the near future, the trend is for the entire population to have access to the Internet, as well as to be present on at least one social networking platform. Regarding companies, it is expected that all, regardless of the business area, have their own website and at least an online social network to disseminate their contents.

A limitation that was found was the fact that, in a universe of 126 companies, only 30 responded to the questionnaire. It did not limit the creation of the strategy, but it would be more interesting to obtain other and varied perspectives on how companies carry out the promotion of their products or services.

Regarding future work, considering that “Alpha” participates in Fairs and Congresses, it would be interesting to implement a strategy dedicated to these specific settings. Bearing in mind that the use of mobile technology has grown substantially, and “Alpha” develops projects covering a wide range of technologies, it would be an asset to create an application in augmented reality. Aimed at potential customers, where it was possible to show in an original and attractive way, the solutions produced, and the services provided by “Alpha”. It would be the individual to control the unfolding of the action, visualizing what interests him best, since Alpha’s audience is very diverse and covers several areas.

In short, this article reveals to be a contribution that confirms the relevance of creating a digital marketing strategy because it enhances the awareness of a brand and the interaction between the company and its target audience.

References

1. Armstrong, G., Kotler, P.: *Marketing: An Introduction* (14th ed.). Pearson (2019)
2. Aukstakalnis, S.: *Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR*. Addison-Wesley Professional (2017)
3. Bartosik-Purgat, M.: Digital marketing communication from the perspective of individual consumers: a cross-country comparison. *Entrepreneurial Bus. Econ. Rev.* 7(3), 205–220 (2019). <https://doi.org/10.15678/eber.2019.070311>
4. Bolotaeva, V., Cata, T.: Marketing opportunities with social networks J. *Internet Soc. Netw. Virtual Communities* 1–8 (2011). <https://doi.org/10.5171/2011.409860>
5. Crook, I., Beare, P.: *Motion Graphics: Principles and Practices from the Ground Up*. Fairchild Books (2016)
6. Ekim, B.: A video projection mapping conceptual design and application: Yekpare. *Turk. Online J. Des. Art Commun.* 1(1), 10–19 (2011). <https://doi.org/10.7456/10101100/002>
7. Guerra, H.S.: Marketing: una visión general y su proceso. In: Velásquez, M.O. (ed.) *Marketing: conceptos y aplicaciones*, pp. 3–37. Universidad del Norte (2014)
8. Hootsuite.: *Digital 2020: Global Digital Overview* (2020). Retrieved January, 2021, from https://p.widencdn.net/1zybur/Digital2020Global_Report_en
9. Jenkins, H.: *Convergence Culture: Where Old and New Media Collide*. NYU Press (2008)
10. Krasner, J.: *Motion Graphic Design: Applied History and Aesthetics* (3rd ed.). Focal Press (2017)
11. Lamb, C.W., Hair, J.F., McDaniel, C.: *Marketing* (11th ed.). Cengage Learning (2011)
12. Leiser, W.: *The World of Digital Art*. H.F. Ullmann (2010)
13. Martínez-Rodrigo, E., Sánchez-Martín, L.: Comunicación y Redes Sociales Presentación. *Rev. ICONO14 Rev. Científica De Comunicación Y Tecnologías Emergentes* 13(2), 1–5 (2015). <https://doi.org/10.7195/ri14.v13i2.887>

14. Patrutiu Baltes, L.: Content marketing - the fundamental tool of digital marketing. Bull. Transilvania Univ. Braşov Ser. V Econ. Sci. **8**(57), 111–118 (2015). http://webbut.unitbv.ro/BU2015/Series%20V/BILETIN%20I/15_Patrutiu.pdf
15. Patrutiu Baltes, L.: Inbound marketing—the most important digital marketing strategy. Bull. Transilvania Univ. Braşov Ser. V: Econ. Sci. **9**(58), 61–68 (2016). http://webbut.unitbv.ro/BU2016/Series%20V/BULETIN%20I/07_Patrutiu_Baltes.pdf
16. Rust, R.T.: The future of marketing. Int. J. Res. Mark. **37**(1), 15–26 (2020). <https://doi.org/10.1016/j.ijresmar.2019.08.002>
17. Ryan, D.: Understanding Digital Marketing: Marketing Strategies for Engaging the Digital Generation (4th ed.). Kogan Page (2017)
18. Schmalstieg, D., Höllerer, T.: Augmented Reality: Principles and Practice. Addison-Wesley Professional (2016)
19. Scholz, J., Smith, A.N.: Augmented reality: designing immersive experiences that maximize consumer engagement. Bus. Horiz. **59**(2), 149–161 (2016). <https://doi.org/10.1016/j.bushor.2015.10.003>
20. Scolari, C.A.: Narrativas transmedia: Cuando todos los medios cuentan. Deusto (2013)
21. Wind, J., Mahajan, V.: Digital marketing. Symphonya. Emerg. Issues Manag. **1**, 43–54 (2002). <https://doi.org/10.4468/2002.1.04wind.mahajan>

Non-invasive Individual Sensing System for Collecting Biometric Indicators



Francisco Monteiro, José Martins, Ramiro Gonçalves, and Frederico Branco

Abstract The daily stress we, as population, are subjected has negative impacts on physical and mental health, making people age faster. With the increase in life expectancy, the population started to search for ways to ensure they reach the older ages with a healthier and better lifestyle, resulting in an increase of interest and demand for thermalism and wellness tourism. Although thermalism is recognized to have several benefits with regard to the combat and prevention of certain pathologies, this has not yet been fully proven. So, in this paper, we presented a sensing system, capable of not only recording thermalist's biometric data, but also getting the data ready and prepared to be processed by any analytic tool with the aim of showing the impact that this sector has on the health of its practitioners.

Keywords Thermalism · Non-invasive sensing system · Wearables · Health and wellness tourism

1 Introduction

The world's population is increasingly living a hectic and stressful life that translates into a decrease in health, both physical and mental, and this is a topic that is growing

F. Monteiro · R. Gonçalves (✉) · F. Branco
Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
e-mail: ramiro@utad.pt

F. Branco
e-mail: fbranco@utad.pt

J. Martins · R. Gonçalves · F. Branco
INESC TEC, Porto, Portugal
e-mail: jose.martins@aquavalor.pt

J. Martins · R. Gonçalves
AquaValor—Centro de Valorização e Transferência de Tecnologia da Água, Chaves, Portugal

J. Martins
Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança, Portugal

awareness of the population regarding the individual lifestyle and has been increasing interest and demand for activities that promote healthy lifestyles [1].

Wellness tourism is a way of merging leisure time with its therapeutic benefits and has been confirmed as one of the sectors that have grown the most in recent years, and so you are seeing an effort from various number of countries starting to invest in these kinds of treatments. And because of this, the spa and wellness tourism sector has been growing in recent times, and so it is expected that it will continue in the coming years as this has been increasingly seen as a method of preventing and mitigating the effect of various pathologies and promoting healthier lifestyles [2–4]. Thermalism is a form of wellness tourism and is defined as the use of natural waters at various temperatures for medicinal use [5].

However, since it is an activity still in development, there is a need to prove the efficiency of treatments and their benefits for its practitioners [6]. So, in this paper, it is presented a sensing system, capable of gathering biometric data from users of thermal spas using a non-invasive device to provide useful information that can prove the benefits of thermalism but also a platform for data visualization even though in a raw stage. The system is supported by a robust architecture designed for easy integration with external systems using the latest technologies.

This paper starts with a brief description of the context and the problem at hand. In the next section, we present the conceptual framework from which the entire system was implemented. The third section is used to detail the material we have at our disposal and is explained the technological choices that were made to develop the system. Next up, it begins with the logical definition that will support the system followed by the actual implementation. Finally, in the last chapter, a final retrospective of the project is made and is presented the limitations that were found.

2 Non-invasive Sensing System in the Wellness Sector

Well-being is comprised of the harmony of the body, mind, and spirit, in which the fundamental elements for reaching the maximum exponent of personal well-being are self-responsibility, physical fitness, relaxation, and mental activity [3]. Over recent times, activities related to health and well-being have been recognized around the world, through the existence of different alternatives for people to stay healthy and to treat themselves to ensure their well-being [7, 8]. But as it is a sector that is quite recent and is still in development that is the need to monitor the practitioners of the thermalism and wellness sector. So, what is the best way to keep track of the evolution of a patient's health condition? Yes, to monitor their biometric data.

2.1 Non-invasive Sensing System

In order to monitor data, firstly we need a system capable of collecting that data. But since we are referring to gathering physiological data, that process is not that linear because it is needed almost hospital-level equipment in order collect, for example, the heart rate or blood pressure. We want to evaluate the efficiency of the treatments that are carried out in the spas, so the equipment to be used by the thermalist's must be chosen with great care, since it is needed to be used during the entire treatment period, which can be up to two weeks and so, we need to balance the device's ease to use, the level of comfort that the device provides, and the efficiency of the data that is read by it [9, 10].

2.2 Wearable Sensing System

Non-invasive sensors are suitable for the human user because they can make painless measurements and have no risk of infection; however, they are complex from a conceptual and design point of view [10]. In order not to negatively affect the user's daily routine, the wearable system should be comfortable and has flexible components, small size, chemically inert, non-toxic, and hypoallergenic, and in order to achieve greater acceptance by the potential users of the system, it has to be accessible, easy to use, non-intrusive and interoperable between various computing platforms. The selection of the sensors is something of utmost importance as they must be able to do data collection/reading accurately and they also must be able to connect to the processing unit so that data can be transmitted [11].

2.3 Wearable Devices

The concept of IoT encompasses situations where network connectivity and computing power expand everyday objects, sensors, and items and affects data exchanges without the need for user handling. Of these devices, those that can be used as accessories or can be worn are called wearables [12].

Smartwatches are wristwatches, similar to normal watches and have a comfortable design, which allows them to be worn constantly; however, they can run applications, track the user's activities and monitor distance traveled, heart rate, sleep monitoring, blood oxygen saturation, and skin surface temperature [13].

Smartbands are very similar to a smartwatch in terms of design; however, apart from obtaining physiological data, they do not have as many features, and currently on the market, smartbands are equipped with sensors that are used to detect and store data of the user's activities continuously [14].

The chest strap is a device that can obtain measurements of the user's heart rate and the electrical activity of the heart during exercise and is worn in direct contact with the skin, like smartwatches and smartbands, and are placed just below the chest muscles. Overall, they are more reliable and accurate than the options mentioned above as they are placed close to the heart; however, they are more uncomfortable to wear on a daily basis [15].

“The incorporation of electronic-based technology into clothing has led to the concept of smart textiles” [13]. These textiles are considered wearable health devices as they have integrated electrodes capable of acquiring physiological signals. For example, breathing pattern is a direct indicator to assess the functioning of the human body as an elevated respiratory rate is accompanied by elevated heart rate, which means that the body is under strain [16]. A sensor could be integrated into a t-shirt that allowed the measurement of breathing patterns to assess an individual response to a determined exercise.

3 Logical Project

With this system, we wanted to keep the costs the lowest possible, but still be able to satisfy the minimum needs of a sensing system which allows the effectiveness of the thermal treatments carried out in the institutions to be analyzed. The cost is a very important factor in the selection of the technologies to be used, especially with regard to the hardware. This is because the treatments, being all done in water, besides the devices having to be minimally resistant, should also keep a low cost when they are replaced, so we anticipated that the durability of the wearable devices will be short in time.

3.1 Wearable Device Chosen

The selection of the wearable device was an important step because it may have a substantial impact on the acceptance of the system by potential users. To this end, the choice should fall on a device, that is, accessible, user-friendly, and non-intrusive. To avoid as much as possible the interference with the user's normal routine, the device should be comfortable, small in size, chemically inert, non-toxic, and hypoallergenic.

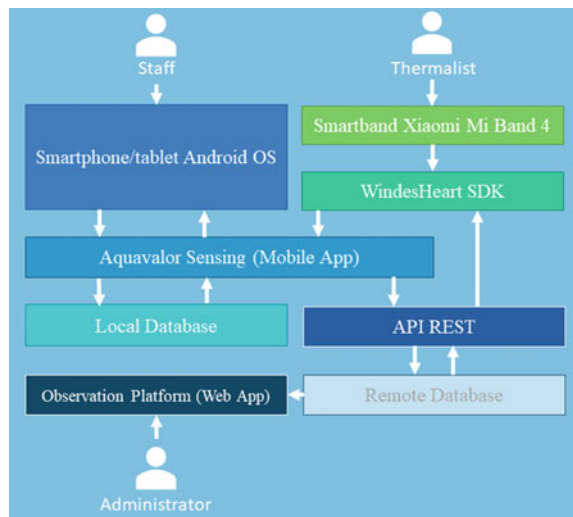
That being said, smartbands were chosen for being capable of collecting the main physiological data at an accessible price, when compared to direct competitors, smartwatches. The fact that these devices (smartbands) have little to no interference in what is the routine of the user was a decisive factor when comparing with other possible choices. Textile-based wearables were still considered; however, the cost of these devices made them impractical for this project. Chest straps were not considered as they were not very practical for the user, given their nature.

3.2 System Architecture

To summarize, the main goal of this project was to implement a non-invasive sensing system capable of collecting thermalist’s biometric measurements during the whole period of treatment and to send it back to a remote database so we could add some layers of data analytics to it in order to analyze and monitor the effects of thermal water treatments on health spa guests. To collect data, we have chosen the Xiaomi Mi Band 4 as the wearable device to be used by all users of the spa because is a cheap one and we have a very limited cost cap. Thus, we started with a conceptual architecture, based on independently loose modules, as we can see in Fig. 1.

The centerpiece of this architecture is the mobile application along with the API REST because it is the link between the data collected and the remote database. Just to explain, there are three actors, the staff, the thermalist, and the administrator. The administrator can be the people responsible to make sure the data we are reading is as expected, as a quality control, and the people that will analyze the information obtained. The thermalist interacts with the system just by using the smartband. At the start of the treatment, the administrator of the spa will register the user in the system (via observation platform) and will assign the smartband. At the end, the thermalism hand it back to the staff. Usually, the treatment takes up to two weeks, and in the meantime, the thermalist uses the smartband. The band has sensors capable of measuring the user’s heart rate, steps taken, and activity level (whether they are awake, in light sleep, or deep sleep) every minute, twenty-four hours a day throughout the treatment period in the spa. This data is stored in the smartband itself. Once the smartband is collected by the staff, he will use the mobile application to connect to the smartband via Bluetooth and gather its data, this will be done with the help of an SDK developed by WindesHeart. The mobile application will be implemented with

Fig. 1 System architecture



Xamarin. After that, the data will be sent to the remote database using the API REST that was developed. The observation platform, a web application built in Outsystems, will then make the data available for viewing.

4 Implementation

As inferred from the system architecture, three modules will be implemented. A mobile application (built-in Xamarin) to collect data from the smartband, an Outsystems web application to manage the spa users, smartband and visualize the data, and a REST API to expose the system’s database. But all these modules need to have some sort of support and follow a series of guidelines in order to achieve an easy integration between them. So, in the E-R diagram below we can see the skeleton of the database that was implemented (Fig. 2).

Exposing this database to the outside modules was developed an API REST, using the Outsystems technology. This API is responsible both for the communication between the database and external applications and for the pre-analysis and preprocessing of the data to be inserted in the database. This data processing occurs when the data is received. Initially, data is stored in temporary tables where analysis is performed. Duplicate and incoherent data is removed for later storage of only the clean data. It was decided to integrate an API for this communication so that the various applications of the system would not have direct access to the database, thus adding an additional level of abstraction and security. Communication is done through HTTP methods, namely GET and POST methods so that external applications can store and consult data in the database.

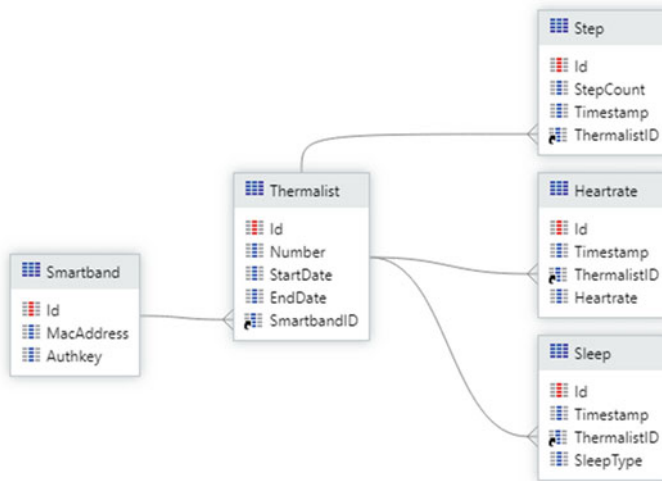


Fig. 2 E-R diagram of the implemented database

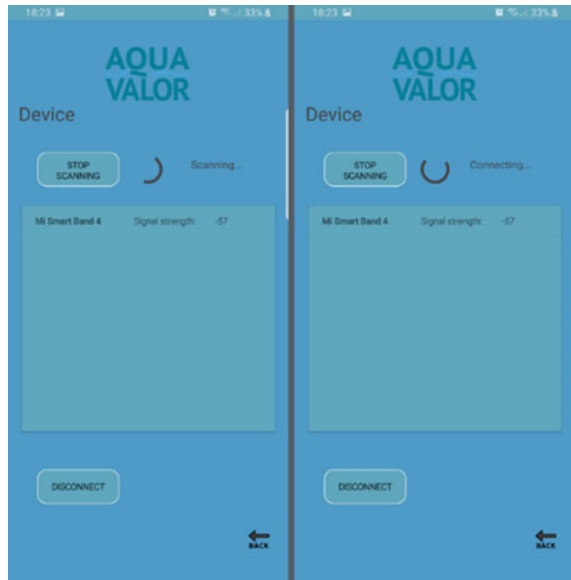
The first tool to consume this API is the SDK, and it is an open-source tool developed by WindesHeart and provides methods to access data that is stored in Xiaomi smartbands. We decided to use it in our system to be the bridge between the smartband and the mobile application. However, for the SDK to be able to communicate with the smartband, a unique authentication key is required for each smartband. To obtain that key, one of the methods provided by the REST API is used, which returns the key corresponding to the MacAddress of the band where the authentication is being made.

In order to the SDK to be used, it has to be triggered by the mobile application. This application works as a link between the data that is collected in the smartband and the remote database and is only used by the staff of the spa. It was developed using Xamarin and is installed on mobile devices (can be smartphone or tablet) with Android operating systems. Using the SDK described above, the application first connects to the band (see Fig. 3) and then gets the data from the smartband and stores it temporarily in a local database, in the mobile device.

This data can be visualized in the application through descriptive graphics (as shown in the figure below) until the moment when the transfer of the data to the remote database begins. The sending of the data is done through HTTP methods, also provided by the REST API. Once the request is complete, the data is removed from the local database and can be seen in the observation platform (Fig. 4).

The visualization/observation platform is mostly used for data quality control purposes and for registering both spa users and smartbands in the system. In this platform, it is possible to see if the data is being correctly stored and if the process of cleaning data is being done properly. The below images illustrates the process of

Fig. 3 Mobile application “Connect device” page. On the left, the scanning for devices, on the right connecting to device



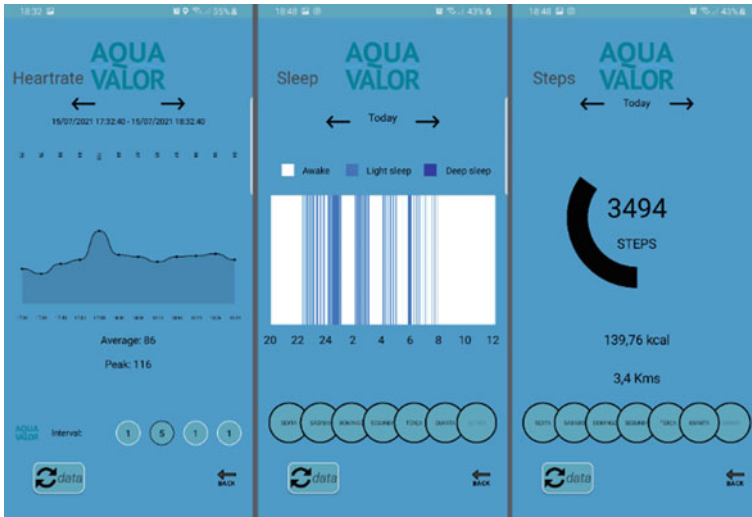


Fig. 4 From left to right

adding a new smartband and a new thermalist to the system. On the left, we have the screen that lists all the smartbands available and the button to add a new one. This data is important to be correct, because otherwise the mobile application will not be able to connect to the smartband. On the right, that is the information of all thermalists registered in the system and the option to add a new one. Each thermalist has an identifier, the start and finish date of the treatment and the smartband that is assigned to him (Fig. 5).

Clicking in the “Add button” will redirect to the respective screen to add a new entry (Fig. 6).

Last but not least, it is the screens that show the data of determined thermalist. Again, this data is still raw, the only processing that was made was just cleaning duplicates and incoherent data, but it is still a good info to have, and we can draw the first conclusions about the treatment (Fig. 7).

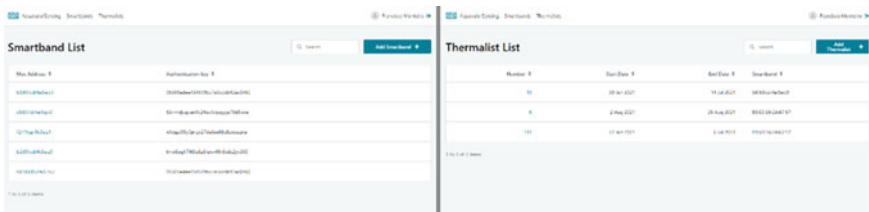


Fig. 5 On the left, the screen that shows the list of smartbands in the system. On the right, the screen that shows the list of thermalists in the system

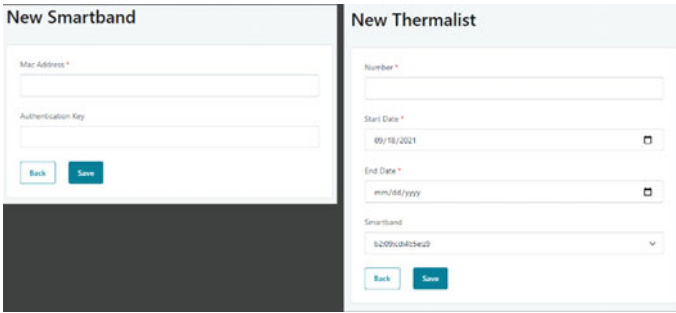


Fig. 6 Screens to add a new smartband (left) and a new thermalist (right)

The screenshot shows a web interface with a navigation bar containing 'AqualorSensing', 'Smartbands', 'Thermalists', and 'Data'. The user's name 'Francisco Monteiro' is visible in the top right. The main section is titled 'Data' and features a 'Thermalist' dropdown menu set to '10'. Below this is a table with three columns: 'HeartRate', 'Steps', and 'Sleep'. The table contains eight rows of data for the date '2 Aug 2021' between 11:00 and 11:06.

HeartRate	Steps	Sleep
66	4	awake
83	4	awake
77	3	awake
74	0	awake
76	4	awake
67	5	awake
70	1	awake
77	5	awake

Fig. 7 Screen to present the raw data that was collected by the sensing system

5 Conclusion

The thermalism sector, despite being growing and seen as a way of promoting health and well-being, still needs to show concrete results. With this in mind, a sensing system capable of obtaining physiological data from thermalists was implemented. This data gives information about the user’s activity level, heart rate, and sleep quality. After analyzing this data, it is possible to infer the impact that treatments based on natural mineral waters have on the health of practitioners in the sector.

5.1 Study Limitations

Although wearables are currently popular devices, they still carry with them somewhat high costs, when used on a large scale. The major limitation of this project was to keep the budget as low as possible. This limited the options when choosing the wearable device to be used. As we proceed with a cheap smartband, the amount of biometric data that could be collected was limited and the precision of this data was not the ideal, so this limited the work that could have been done.

5.2 Future Work

With this project, it was possible to show a good proof of concept regarding the implementation of a non-invasive sensing system. The implementation of such a system, besides being able to bring important data that can prove the benefits of thermalism, also serves to control the treatments that are carried out in spas and thermal baths. Still, it is to improve the quality of the data in the system.

One way to improve the information obtained would be to use the smartbands in a period prior to the treatment to have a good basis for comparison and at the same time increase the amount of data present in the knowledge bases. The use of devices with a higher degree of confidence and which make other types of measurements available would also be an added value, since it would increase the diversity of data that could be obtained, as well as increasing their accuracy, with an impact on the richness of the analysis.

Since only the cleaning was done for a better visualization of the data, it is still possible to add an additional layer of data analytics that can translate data into useful information. The idea would be to use tools such as Power BI to produce treatment reports and present them through interactive dashboards, thus providing insights for both those who evaluate treatments and for the spa users themselves to have an idea of their evolution.

The modularity of the architecture implemented in the system opens the door to integration with external systems. For example, the fact that the database has its own API enables integration with analytical systems, as in the previous example of Power BI. The web application itself, which oversees the management of the data, the thermalists, and the smartbands, can easily be extended to a management system of the spa itself, where the employees, the spa occupancy, etc., can be managed. All this is done in a simple “plug-and-play” way.

5.3 Final Thoughts

In general, the final balance of the project is positive. We managed to implement what was idealized and we are satisfied with the final result obtained. Although it is not similar to an individual medical analysis, with a bit of data analysis on top of the data, we can obtain through the sensing system will undoubtedly serve as a study basis to prove the effects that thermalism has on the health of its practitioners.

References

1. Neto, R.R.: PANDEMIA TRAZ NOVO DESTAQUE À SAÚDE E BEM-ESTAR. Available: <https://www.publituris.pt/2021/04/09/pandemia-traz-novo-destaque-a-saude-e-bem-estar/>, Accessed
2. Mesa, B.C.B.: Saúde e bem-estar pela água: Turismo do Porto e Norte aposta na internacionalização das termas, Available: <https://boacamaboamesa.expresso.pt/boa-vida/2020-12-28-Saude-e-bem-estar-pela-agua-Turismo-do-Porto-e-Norte-aposta-na-internacionalizacao-das-termas>, Accessed
3. Koncul, N.J.E.r.-E.i.: Wellness: a new mode of tourism **25**, 525–534 (2012)
4. Lee, S.-R., Lee, Y.-L., Lee, S.-Y.J.J.o.T.K.S.o.I.M.: Short-term effect of gyeongsangnam-do wellness tourism on inflammatory and immune system markers, emotion, fatigue, stress, quality of life, and quality of sleep **9**, 19–35 (2021)
5. Erfurt-Cooper, P., Cooper, M.: Health and Wellness Tourism: Spas and Hot Springs. Channel View Publications (2009)
6. Mairós, O.: Chaves aposta forte no turismo termal. Available: <https://rr.sapo.pt/noticia/pais/2018/05/07/chaves-aposta-forte-no-turismo-termal/112559/>, Accessed
7. Azman, I., Chan, K.J.H.: Wellness, tourism: healthy tourists, h.b.: health and spa tourism business: tourists' profiles and motivational factors **9**, 24 (2010)
8. Utama, I., Nyandra, M.J.I.G.B.R.U., Health, M.N., Types, W.T.I.: Development potentials In: Bali, I.I.J.O.S.S., Studies, E.R.: Health and Wellness Tourism Industry: Types and Development Potentials in Bali, Indonesia. 1, 08-12 (2021)
9. Majumder, S., Mondal, T., Deen, M.J.J.S.: Wearable sensors for remote health monitoring **17**,130 (2017)
10. Lmberis, A., Dittmar, A.J.I.E.i.M., Magazine, B.: Advanced wearable health systems and applications-research and development efforts in the European union **26**, 29–33 (2007)
11. Al-Khafajiy, M., Baker, T., Chalmers, C., Asim, M., Kolivand, H., Fahim, M., Waraich, A.J.M.T.: Applications: remote health monitoring of elderly through wearable sensors. **78**, 24681–24706 (2019)
12. Metcalf, D., Milliard, S.T., Gomez, M., Schwartz, M.J.I.p.: Wearables and the internet of things for health: wearable, interconnected devices promise more efficient and comprehensive health care **7**, 35–39 (2016)
13. Dias, D., Paulo Silva Cunha, J.J.S.: Wearable health devices—vital sign monitoring, systems and technologies **18**, 2414 (2018)
14. Knighten, J., McMillan, S., Chambers, T., Payton, J.: Recognizing social gestures with a wrist-worn smartband. In: 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), pp. 544–549. IEEE (2015)
15. Cosoli, G., Spinsante, S., Scalise, L.J.M.: Wrist-worn and chest-strap wearable devices: systematic review on accuracy and metrological characteristics 107789 (2020)
16. Coyle, S., Morris, D., Lau, K.-T., Diamond, D., Moyna, N.: Textile-based wearable sensors for assisting sports performance. In: 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks, pp. 307–311. IEEE (2009)

Machine Learning Applied to Marketing

Image Processing: Impact of Train and Test Sizes on Custom Image Recognition Algorithms



Luis Marques, Luca Lopes, Miguel Ferreira, Cristina Wanzeller, Pedro Martins, and Maryam Abbasi

Abstract This paper intends to demonstrate results on applying machine learning algorithms to process image recognition to identify professions. This kind of project points us to a relation between humans and machines, so in a way, we might say that the human brain and vision process are being passed to a machine in order to bring us many benefits in our daily life. In this paper, we decided to compare how different parameters influence the performance and accuracy of the following neural networks: EfficientNetB0, NASNetMobile, MobileNetV2, ResNet50, InceptionV3, and DenseNet121.

Keywords Machine learning · Image recognition · ResNet · InceptionV3 · MobileNetV2 · DenseNet · NASNetMobile · EfficientNet

L. Marques · L. Lopes · M. Ferreira
Polytechnic of Viseu, Viseu, Portugal
e-mail: estgv5989@alunos.estgv.ipv.pt

L. Lopes
e-mail: stgv13082@alunos.estgv.ipv.pt

M. Ferreira
e-mail: estgv7447@alunos.estgv.ipv.pt

C. Wanzeller · P. Martins (✉)
CISeD—Research Centre in Digital Services, Polytechnic of Viseu, Viseu, Portugal
e-mail: pedromom@estgv.ipv.pt

C. Wanzeller
e-mail: cwanzeller@estgv.ipv.pt

M. Abbasi
CISUC—Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal
e-mail: maryam@dei.uc.pt

1 Introduction

Image classification is used in various applications, such as security, educational, and promotional systems. In recent years, much research has been done to design automated systems to extract fundamental features from images. Convolutional neural network (CNN) is an effective method for image classification which uses convolutional, pooling, and fully connected layers for the learning process [11]. CNN has redefined the state of the art in many real-world applications, such as facial recognition, image classification, human pose estimation, and semantic segmentation [19]

The training of machine learning models for image processing is a process that consumes many resources [10]. Despite the computational power we currently have at our disposal in personal computers, especially in graphics processing units (GPU) [5], this power is still limited for usage in computer vision, mainly in CNNs. This class of application usually relies on heavy computations on massive datasets. Therefore, parallel computing is traditionally considered to run the training process in a feasible time using the GPU [2]. Acquiring these hardware implies risks under and overutilization, depreciation of the hardware, and failures. There are also costs related to maintenance, energy, and human resources [2]. However, there are solutions in the cloud in all significant providers, Amazon AWS [16], Microsoft Azure [1], and Google Cloud [3], that can aid in this process, making the machine learning process more accessible. The researcher/developer does not take the risk of acquiring hardware that can quickly become obsolete and only pay for the resources that he will use [10]. However, some free alternatives in the cloud, namely Google Colaboratory (commonly referred to as “Google Colab” or just “Colab”) [6] and Paperspace Gradient [14], are the ones used in this work. Computational resources are not the only requirement for training machine learning models; in-depth knowledge of applied mathematics and deep learning libraries is also required [10]. While this might pose a problem, there are, however, libraries that simplify this challenge, namely ImageAI. ImageAI is a Python library built to empower developers, researchers, and students to build applications and systems with self-contained deep learning and computer vision capabilities using simple and few lines of code [13].

This study aims to compare some of the ImageAI provided algorithms ResNet50, MobileNetV2, InceptionV3, DenseNet121, and two more EfficientNetB0, NASNet-Mobile, which were made available through some custom code, on the IdenProf [12] dataset and two customizations of it, with different train and test sizes in order to recognize professions in images, using Paperspace Gradient and Google Colab as research environments. This document is organized into five sections. In Sect. 2, we discuss related work in the area. Section 3 presents the method applied in study. Section 4 shows the obtained results. Finally, in Sect. 5, conclusions are drawn, followed by future work guidelines.

2 Related Work(s)

This paper focuses on the topic of computer vision and its technology. It is easy for humans to describe and understand the objects that we see from the world. Our visual system can perceive a three-dimensional structure with enough information, such as the objects' shape, appearance, and color. However, this is not easy for a computer [20]. Researchers in this field try to mimic the capacity of how human vision works using computers. However, it is not an easy task, and literature on artificial visual processing is usually categorized into visual processing algorithms, which consist in the recreations of the human vision, and classifiers, which are remodeling of the human decision techniques [4].

Computer vision is a vast research field where mathematics, geometry, and physics are applied [20]. However, some tasks are commonly accomplished with computer vision, object detection, recognition, and classification. This paper is focused on the classification part. Image classification was once a task that required domain expertise and the use of problem-specific models. Much of this has changed with the emergence of deep learning as a general-purpose modeling technique for predictive tasks in computer vision. Both the machine learning literature and image classification contests are now dominated by deep learning models that often do not require domain expertise since such models identify and extract features automatically, eliminating the need for feature engineering [9].

Usage of libraries like ImageAI allows us to train and generate image classification models with CNN without extensive knowledge of the inner network workings. However, it allows us to use its potential as this kind of algorithm is broadly used in computer vision. However, the requirements of computation for training models with these algorithms are high. With our experimentation, we used ResNet50 [7], DenseNet121 [8], and InceptionV3 [17] as these are also used by the majority of studies related to image object recognition, and we also used MobileNetv2 [15], NASNetMobile [21], and EfficientNetB0 [18]. CNNs provide high accuracy. The main reason for this is because the number of features increases dramatically. The research done about computer vision relies on the accuracy of the validation and improving that accuracy. Also, studies are showing us that if we keep training our model with thousands of pictures, we can reach into overfitting issues. There is a balance that needs to be respected when training models [10] Last but not least, as shown by some studies, the quality of the datasets impacts creating and training a model.

Setting up an on-premises solution to research and build models from machine learning algorithms can be expensive and not the only available option. All major cloud providers offer services in the cloud for the same purpose with access to a huge computing capacity. Some of these offers are Amazon Sagemaker [16], Microsoft Azure Machine Learning [1], Google Cloud AI infrastructure [3], and also free options like Google Colab [6] or Paperspace Gradient [14]. These last two platforms used in the experiment.

3 Experimental Setup

This project's architecture consists of using Paperspace Gradient provided Docker containers, which provided the necessary infrastructure for the code developed in Jupyter Notebooks and the base storage. Google Colab was also used (and integrated with Google Drive) in the final work for collecting graphics and metrics of TensorBoard logs.

The Paperspace Gradient free supplied containers include dedicated NVIDIA Maxwell GPU with 8 GB of GPU memory, 30 GB of memory, 8vCPU, and 5 GB of storage space. These resources are all free but with a limit of the run session of 6 h maximum. For Google Colab, the type of GPUs available in Colab varies over time, often including Nvidia K80s, T4s, P4s, and P100s. The standard available RAM in Colab is 12 GB. It should be noted that these resources are shared between users of the platform, so the available capacity of the resources varies over time.

3.1 Datasets

The author of ImageAI [13] Python library also created a dataset IdenProf [12], which was used as the base for this study, but we also created two custom datasets from it. The IdenProf dataset contains 11,000 images that span over ten categories of professions. Each profession category consists of 1100 images, 900 of which are used for training and the remaining 200 for testing. Our custom datasets consist of the same 1100 as the base dataset. However, the training and test sizes are different, 800 and 300 for one dataset and 1000 and 100. These datasets will be referenced as DS100, DS200, and DS300, matching the respective test sizes from now on.

The images in the dataset have a resolution of 224 * 224 pixels and represent subjects dressed in uniforms of their respective professions. The dataset distribution as of represented subject is as follows, 19.4% female an 80.6% male, 91.1% white, and 8.9% dark skins.

The process of acquisitions of the dataset images, as the dataset author describes it *“The images in the dataset were obtained from Google Image search. The images were searched and collected based on the 15 most populated countries in the world. The dataset does not comply with EU GDPR has the individuals whose images were contained were not explicitly contacted for consent”* [12].

3.2 Parameters

ImageAI library provides several algorithms that can be used for image classification, namely ResNet50, DenseNet121, MobileNetV2, and InceptionV2. Other two

Table 1 Setup parameters

Training cycle						
Epochs 25						
Per Epoch iterations						
	DS100		DS200		DS300	
Batch size	Train	Validation	Train	Validation	Train	Validation
16	625	62	562	125	500	187
32	312	31	281	62	250	93
64	156	15	140	31	125	46

algorithms not present in ImageAI were also used, namely NASNetMobile and EfficientNetB0. All algorithms were used to train models from the datasets during 25 epochs in three different image batch sizes 16, 32, and 64 for the three datasets.

A preliminary test using a batch size of 128 was additionally thought, but due to resources limitations and time constraints was not carried over. In Table 1, we can observe a resume of setup parameters.

4 Results and Analysis

From the facts gathered, accordingly to Table 2 InceptionV3 was the algorithm that lost less trainable parameters followed by ResNet50, all other algorithms had losses over 1%. Both algorithms also detect more parameters, being, in this case, ResNet50 has more parameters detected, which translates into more significant model sizes. MobileNetV2 algorithm was the one that lost most parameters that could be trained; also, this was the one that detected fewer parameters, which translates to smaller size models. A final fact we can observe from analyzing all algorithms, as the number of parameters detected grows, so does the model sizes grow proportionally.

As we can observe in Table 3, related to temporal data, the faster algorithm in dataset DS100 was MobileNetV2 for batch sizes 16 and 64; however, for a batch size of 32, it was EfficientNetB0. For dataset DS200, MobileNetV2 was faster for batch size 16, in batch size 32, it was EfficientNetB0, and in batch size 64, it was NASNetMobile. For dataset, DS300 EfficientNetB0 was faster in batch size 16 but slowest in batch size 32, where InceptionV3 was faster with batch size 64.

In other facts gathered, for dataset DS100, there was a reduction in the train time from a batch size of 16 to a batch size of 64. As of dataset DS200, that only happened with NASNetMobile and EfficientNetB0, from batch size 16 to batch size 32, there was an overall reduction excluding MobileNetV2. In dataset DS300, only NASNetMobile had time reduction from batch size 16 to batch size 64, and all others

Table 2 Algorithms facts collected

	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
Total pa rams	23,608,202	7,047,754	21,823,274	2,270,794	4,280,286	4,062,381
Trainable pa rams	23,555,082	6,964,106	21,788,842	2,236,682	4,243,548	4,020,358
Non trainable pa rams	53,120	83,648	34,432	34,112	36,738	42,023
Of params model (%)	0.23	1.19	0.16	1.50	0.86	1.03
Size (MB)	90.5	27.8	83.9	9	17.9	16

Table 3 Training times facts

Dataset	Batch size	ResNet50	DenseNet121	InceptionV3Mobile	MobileNetV2	NASNetMobile	EfficientNetB0
DS1000	16	52m46s	48m07s	47m50s	46m59s	58m38s	47m22s
	32	55m46s	50m21s	50m27s	56m00s	51m24s	50m11s
	64	43m09s	42m49s	43m06s	41m46s	47m45s	42m49s
DS2000	16	40m5s	41m23s	42m11s	38m4s	40m5s	51m11s
	32	36m43s	36m38s	35m53s	39m45s	40m32s	36m17s
	64	54m02s	47m23s	46m45s	50m43s	46m22s	46m20s
DS200	16	50m01s	42m56s	45m37s	53m13s	52m54s	40m13s
	32	48m55s	46m15s	45m58s	46m46ss	47m11s	52m48s
	64	53m26s	51m11s	47m01s	53m32s	48m04s	47m48s

had time increased. For DenseNet121 and InceptionV3, as the batch size increased, so did the train time. We can extract from the facts that the best training time obtained across all datasets and all batch sizes was for InceptionV3 in dataset DS200 and batch size of 32.

According to Tables 4 and 5, we can observe that for all datasets and all batch sizes, the training accuracy is slightly better than validation accuracy; however, for NASNetMobile algorithm only for batch sizes of 16, this difference is slight for other batch sizes, validation accuracy is much lower, and the same behavior occurs with MobileNetV2 for dataset DS100 on a batch size of 32 and remaining datasets in batch size 64. Also for an expected behavior with all algorithms across all datasets and batch sizes, accuracy tends to increase with batch size increase, as for validation accuracy that tendency is inverse, meaning with batch size increase validation accuracy tends to decrease, but for some batch sizes and algorithms that is not always the case. Analyzing Table 4, another result was gathered, for all algorithms without exception, in dataset DS100 higher accuracies were achieved than those of DS200, and these were greater than DS300.

Table 6 is relative to the training loss. This is a metric worth analyzing, as this can indicate how good the predictive model is, as lower the loss, the better the predictions are. Observing these values, we see some similarities between some algorithms. In datasets, DS200 and DS300, all algorithms excluding EfficientNetB0 decreased loss as the batch size increased, as for EfficientNetB0 decreased from batch size 16–32 and increased slightly from batch size 32–64, still lower than batch size 16. EfficientNetB0 and MobilNetV2 for dataset DS100 decreased loss as the batch size increased, while all other algorithms had the same behavior, decreasing from batch size 16–32 and increasing from 32 to 64. NASNetMobile had the lower loss in all datasets, in DS100 was in batch size 32, as for the others was with 64.

Observing the values in Table 7, relative to the validation loss, we saw some similarities between some algorithms, and MobileNetV2 and NASNetMobile stand out as having much higher losses than the rest in all datasets for the majority of batch sizes. For EfficientNetB0, the loss behavior was the same per batch size across the datasets, increasing loss as the batch size increased. ResNet50 for dataset DS100 and DS200 performed the same with loss decrease from batch size 16–32, but with an increase in 64, while for dataset DS300, as the batch size increased, the loss decreased. DenseNet121 and InceptionV3 had the same behavior as ResNet50 in dataset DS100, while in dataset DS200, InceptionV3 maintained the same behavior, and DenseNet121 increased loss as batch size increased for DS200 and DS300 datasets. The lower loss was obtained with DenseNet121 for dataset DS100 and batch size of 64, while for the remaining datasets, InceptionV3 had a lower loss with batch size 32.

Table 8 represents the difference between the losses from training and from validation. This is a fundamental metric to pay attention to, as we can check if the models trained might be overfitting or underfitting. Ideally, this difference should be zero, or as close as we could get, but usually, some overfitting occurs. At first glance, NASNetMobile stands out as having a more significant difference than the

Table 4 Train accuracy on datasets across batch sizes

Dataset	Batch Size	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
DS100	16	0.8293	0.8877	0.867	0.8321	0.918	0.8481
	32	0.898	0.9111	0.9258	0.8471	0.9444	0.8586
	64	0.8392	0.9044	0.9045	0.857	0.9297	0.8669
DS200	16	0.8251	0.8659	0.8575	0.808	0.9154	0.8231
	32	0.8695	0.8884	0.898	0.8388	0.9296	0.8553
	64	0.8704	0.9041	0.926	0.8499	0.9335	0.8494
DS300	16	0.7979	0.8524	0.8356	0.8026	0.9145	0.8155
	32	0.8129	0.871	0.8845	0.8285	0.9125	0.8395
	64	0.8696	0.8835	0.924	0.8436	0.9381	0.8444

Table 5 Validation accuracy across batch sizes and datasets

Dataset	Batch Size	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
DS100	16	0.8034	0.8367	0.8286	0.8105	0.8024	0.8054
	32	0.8115	0.8313	0.8406	0.1	0.1917	0.7927
	64	0.7853	0.8327	0.8468	0.8075	0.2601	0.7923
DS200	16	0.7845	0.821	0.814	0.7835	0.7055	0.784
	32	0.7974	0.8196	0.8311	0.7263	0.2041	0.7823
	64	0.7918	0.814	0.8191	0.1003	0.126	0.7777
DS300	16	0.7784	0.8142	0.8031	0.7878	0.6397	0.7841
	32	0.7779	0.8095	0.8222	0.4694	0.2023	0.7769
	64	0.7952	0.8077	0.8173	0.1005	0.1345	0.7748

Table 6 Train loss across batch sizes and datasets

Dataset	Batch Size	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
DS100	16	0.4973	0.3427	0.3834	0.4956	0.2385	0.4351
	32	0.3066	0.2774	0.224	0.4427	0.1705	0.4128
	64	0.4583	0.2882	0.2749	0.4227	0.2075	0.3772
DS200	16	0.5881	0.3981	0.415	0.5556	0.2527	0.5101
	32	0.373	0.3325	0.2943	0.4764	0.2156	0.4117
	64	0.3666	0.2879	0.2195	0.4316	0.199	0.432
DS300	16	0.5814	0.438	0.4736	0.5734	0.261	0.5373
	32	0.5338	0.3879	0.3352	0.5006	0.2491	0.4654
	64	0.3784	0.3397	0.2248	0.4619	0.18	0.4504

Table 7 Validation loss across batch sizes and datasets

Dataset	Batch Size	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
DS100	16	0.6164	0.5086	0.5151	0.5727	0.6552	0.6203
	32	0.5803	0.5332	0.589	2.924	4.907	0.6632
	64	0.6683	0.4919	0.5327	0.6025	3.721	0.6832
DS200	16	0.6401	0.5507	0.5491	0.6252	0.9749	0.6499
	32	0.6386	0.5514	0.5429	0.7716	3.91	0.6525
	64	0.672	0.544	0.6154	3.227	4.247	0.6919
DS300	16	0.6501	0.5431	0.5533	0.62	1.179	0.636
	32	0.6469	0.5497	0.5378	1.632	3.329	0.6439
	64	0.6239	0.5538	0.5835	3.524	4.561	0.6943

Table 8 Difference between validation loss and training loss across batch sizes and datasets

Dataset	Batch Size	ResNet50	DenseNet121	InceptionV3	MobileNetV2	NASNetMobile	EfficientNetB0
DS100	16	0.1191	0.1659	0.1317	0.0771	0.4167	0.1852
	32	0.2737	0.2558	0.365	2.4813	4.7365	0.2504
	64	0.21	0.2037	0.2578	0.1798	3.5135	0.306
DS200	16	0.052	0.1526	0.1341	0.0696	0.7222	0.1398
	32	0.2656	0.2189	0.2486	0.2952	3.6944	0.2408
	64	0.3054	0.2561	0.3959	2.7954	4.048	0.2599
DS300	16	0.0687	0.1051	0.0797	0.0466	0.918	0.0987
	32	0.1131	0.1618	0.2026	1.1314	3.0799	0.1785
	64	0.2455	0.2141	0.3587	3.0621	4.381	0.2439

rest for all datasets, which indicates a case of overfitting. MobileNetV2 also generated models overfitting in all datasets, for batch size 32 in DS100 and batch sizes 32 and 64 for DS300; as for batch size 16 in all datasets, this algorithm trained some of the best-adjusted models. Overhaul as the batch size increases, so do the generated models overfitting. The less overfitting models for datasets DS100 and DS300 were MobileNetV2 with a batch size of 16, as for DS200, it was ResNet50 also with a batch size of 16.

5 Conclusions and Future Work

After testing the six algorithms and observing the results, we can conclude that all algorithms tend to achieve higher accuracy when the dataset train size increased and test size increased. Another conclusion we extract is that accuracy is achieved between algorithms and batch sizes are similar, with slight differences. The best algorithm does not exist; they all tend to adapt to the datasets.

For the three datasets, the algorithm which achieved higher accuracy under 25 epochs was InceptionV3. Excluding NASNetMobile and MobileNetV2, which presented overfitting, all the other algorithms had validation/training loss differences lower, meaning slightly less overfitting than InceptionV3. Considering accuracy and validation/training loss difference, DenseNet121 appears to be a better algorithm for the datasets in the study.

EfficientNetB0 in all datasets showed that increasing batch size, the accuracy decreased slightly; the same happened to ResNet50 and DenseNet121. This result makes us believe that increasing batch sizes are not beneficial for these algorithms. MobileNetV2, contrary to previous ones, seems to increase accuracy slightly, which seems that increasing batch size might benefit this algorithm. Given its generated smaller model size, this algorithm might be worth exploring in situations where limited resources are available.

As cloud platforms allow developers, scientists, and researchers to use free resources for machine learning, despite its limitations, these solutions prove worth the time exploring, as it would be in-viable carrying out this study on typical household computers. Paperspace Gradient proved us to be a powerful and versatile platform for carrying out these tests.

5.1 Future Work

Our tests showed a slight tendency for more accurate models trained on datasets with more extensive train sets. Another tendency was that bigger batch sizes reduce both train time and accuracy. Similar tests could be done on different datasets and adaptations from the same dataset to validate and improve this study, trying to test bigger batch sizes and more extensive datasets to see if these findings still hold valid.

Acknowledgements “National Funds fund this work through the FCT—Foundation for Science and Technology, IP, within the scope of the project Ref UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD), the Polytechnic of Viseu, for their support.”

References

1. Azure, M.: Azure machine learning—ml as a service—microsoft azure, June 2021. URL: <https://azure.microsoft.com/en-us/services/machine-learning/>. Retrieved: 2021-06-29
2. Carneiro, T., Medeiros Da No’Brega, R. V., Nepomuceno, T., Bian, G.-B., De Albuquerque, V. H. C., Filho, P.P.R.: Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access*, **6**, 61677–61685 (2018)
3. Cloud, G.: Ai infrastructure ml and dl model training—google cloud, June 2021. URL: <https://cloud.google.com/ai-infrastructure>. Retrieved: 2021-06-29
4. Czimmermann, T., Ciuti, G., Milazzo, M., Chiurazzi, M., Roccella, S., Oddo, C.M., Dario, P.: Visual-based defect detection and classification approaches for industrial applications—a survey. *Sensors* **20**(5) (2020)
5. Farber, R.: Chapter 2—cuda for machine learning and optimization. In: *CUDA Application Design and Development* (pages 33–61). Morgan Kaufmann, Boston (2011)
6. Google.: Colab: Frequently asked questions, June 2021. URL: <https://research.google.com/colaboratory/faq.html>. Retrieved: 2021-06-26
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015)
8. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks (2018)
9. Hull, I.: *Machine Learning for Economics and Finance in TensorFlow 2: Deep Learning Models for Research and Industry*, 1st edn. Apress (2021)
10. Martins, M., Mota, D., Morgado, F., Wanzeller, C., Martins, P., Abbasi, M.: Imageai: comparison study on different custom image recognition algorithms. In: Rocha, A., Adeli, H., Dzemyda, G., Moreira, F., Ramalho Correia, A.M. (eds.), *Trends and Applications in Information Systems and Technologies* (pp. 602–610). Springer International Publishing, Cham (2021)
11. Momeny, M., Latif, A.M., Sarram, M.A., Sheikhpour, R., Zhang, Y.D.: A noise robust convolutional neural network for image classification. *Results Eng.* **10**, 100225 (2021)
12. Olafenwa, M.: Idenprof, June 2021. URL: <https://github.com/OlafenwaMoses/IdenProf>. Retrieved: 2021-06-26
13. Olafenwa, M.: Imageai, June 2021. URL: <https://github.com/OlafenwaMoses/ImageAI>. Retrieved: 2021-06-26
14. Paperspace.: Gradient—effortless deep learning at scale, June 2021. URL: <https://docs.paperspace.com/gradient/> Retrieved: 2021-06-26
15. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C.: Mobilenetv2: Inverted residuals and linear bottlenecks (2019)
16. Services, A.W.: Amazon sagemaker—machine learning—amazon web services, June 2021. URL: <https://aws.amazon.com/sagemaker/>. Retrieved: 2021-06-29
17. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision (2015)
18. Tan, M., Le, Q.V.: Efficientnet: Rethinking model scaling for convolutional neural networks (2020)
19. Tran, H.-D., Bak, S., Xiang, W., Johnson, T.T.: Verification of deep convolutional neural networks using imagestars. In: Lahiri, S.K., Wang, C. (eds.) *Computer Aided Verification*, pp. 18–42. Springer International Publishing, Cham (2020)

20. Zhang, Y.B.B.X.M.: Genetic Programming for Image Classification: An Automated Approach to Feature Learning, Volume 24 of Adaptation, Learning, and Optimization, 1 edn. Springer International Publishing (2021)
21. Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V.: Learning transferable architectures for scalable image recognition (2018)

Neuromarketing Technologies

Demand Analysis for the Creation of a Technology Application to Improve Decision Making



Micaela Silva-Barragán and Carlos Ramos-Galarza

Abstract This chapter proposes a mobile application aimed at improving the decision-making process through conscious stimulation of the prefrontal cortex using a smartphone application. It is expected to be useful for adolescents and young adults, considering the brain maturation and important decisions taken at this stage, which can have positive or negative consequences. The research design followed consists of a longitudinal descriptive study. The positive perception and social reception of the prototype should be considered as results. Finally, the demand analysis and its results are presented.

Keywords Decision making · Executive functions · Mobile applications · Frontal lobule

1 Introduction

Adolescence is conceived as a stage of rapid and constant changes that affect the brain and cognitive functions. In addition to adapting to changes, teenagers must acquire new skills that will allow them to function with greater autonomy and independence while developing their own identity. The activities outside the teenagers' core family require their reasoning, cognitive and emotional resources, which will allow them to make decisions with positive or negative consequences [1]. In this sense, making decisions occupies a central place and will be a process that requires cognitive skills typical of the prefrontal cortex (PFC). This brain zone oversees executive functions (FE) and, according to Luria's brain organization model, corresponds to the third functional unit responsible for complex processes such as planning, regulation, and verification of mental and behavioral activity. It should be emphasized that the PFC

M. Silva-Barragán · C. Ramos-Galarza (✉)
Facultad de Psicología, Pontificia Universidad Católica del Ecuador. Av. 12 de Octubre y Roca,
Quito, Ecuador
e-mail: caramos@puce.edu.ec

C. Ramos-Galarza
Centro de Investigación en Mecatrónica y Sistemas Interactivos MIST/Carrera de Psicología,
Universidad Tecnológica Indoamérica. Av. Machala y Sabanilla, Quito, Ecuador

will finish developing around 20 or even 25 years old, resulting in some late and progressive maturity that implies that the adolescent brain is reorganized and more sensitive to environmental experiences [2–4].

Exposure to risky situations, or what is called “risk behavior,” is something that characterizes adolescence and has a close relationship with decision making. It is in this context that a teenager will have to apply their cognitive and emotional resources to choose an option, balancing between their safety and the subjective benefits of taking risks such as new experiences, greater status among peers, satisfaction, and pleasure [5]. This is how the proposed mobile application aims to cognitively stimulate the PFC to boost its development and, consequently, direct the young users to better decision making in different areas of their life. The technology has been proven to be an ally even in cases requiring neuropsychological rehabilitation [6]. Also, it is expected that teenagers will be interested in using technology for this purpose if the application is similar to a video game. It should be emphasized that the adolescent problem around this topic is very complex and such cognitive stimulation should be combined with other psychosocial interventions.

The application of technological tools has not only been useful in cases of neuropsychological rehabilitation after brain damage and cognitive impairment, it should be noted that different interventions of this type have helped to stimulate cognitive functions. For example, in Aldana, García, and Jacobo’s study, it was shown that the use of technological tools such as computer programs and online platforms favors cognitive processes such as attention, memory, intelligence, and language in elderly population [7]. Considering that adolescence is a vital stage with greater neuronal plasticity, it could be enhanced the improvement of cognitive functions that are the basis of the decision-making process. Another situation in which the influence of mobile applications is demonstrated occurs in the study by Salazar and Gallegos, who evidenced that the technological intervention improved several functions, among them the executive function of cognitive flexibility [8]. In a more specific way, Schwartz conducted a review of how the integration of technology can promote the academic performance in students. This is achieved by stimulating different executive functions such as working memory, planification and time management [9]. At this point, it is important to remember that FE have a crucial role in the decision-making process.

Therefore, this chapter will briefly explain the technological proposal and then analyze the demand for it in a group of young people between 17 and 22 years old.

2 Technological Proposal for Decision Making

The technological proposal is a video game for mobile devices, due to their daily use by teenagers. Within the application, the users must register to create a character of their choice. Afterward, this can be customized with different clothing, hair, or skin color. Once this is done, six themes appear on the screen: party, friends, courtship, drugs and alcohol, school, and family. These themes were chosen because

suit adolescents' daily life and are also of interest to them. Moreover, each theme includes decision-making situations, which are displayed graphically and in writing to provide an explanation of what is happening and the options to choose. These decisions respond to the presentation of a dilemma or problem in which the character must choose which path to take. Finally, the users are shown the consequence (positive or negative) of their decision and feedback is offered about the risks presented in that scenario and the importance of making decisions that are projected to positive consequences. This proposal is explained in more detail in a previous article.

As mentioned above, the mobile application will allow the users to consciously exercise the decision-making process while offering an alternative with ecological validity where the situations presented are usual at this life stage. While other articles have already highlighted the importance of using technology to develop and improve brain processes, a mobile application is more accessible to the users [10]. Also, it has been shown that the use of technology contributes significantly even to the neuropsychological treatment of patients with acquired brain damage or in the stimulation of multiple intelligences such as linguistic-verbal and logical-mathematical [11, 12]. In this sense, this technological proposal aims to improve the brain decision-making process, which is highly important for adolescents and young adults to unfold in society.

As mentioned above, the mobile application will allow the users to consciously exercise the decision-making process while offering an alternative with ecological validity where the situations presented are usual at this life stage. While other articles have already highlighted the importance of using technology to develop and improve brain processes, a mobile application is more accessible to the users [10]. Also, it has been shown that the use of technology contributes significantly even to the neuropsychological treatment of patients with acquired brain damage or in the stimulation of multiple intelligences such as linguistic-verbal and logical-mathematical [11, 12]. In other investigation, Rossignoli-Palomeque demonstrates how an informatics application called NEXXO can improve the attentional control, executive control, and self-regulation in children with TDAH [13]. The use of real-life situations in the mobile application is a strategy that speeds up the acquisition of experience and allows the user to manipulate a complex circumstance without being directly exposed. In that way, it is possible to improve the necessary apprenticeship of a good decision-making process [8]. In this sense, this technological proposal aims to improve the brain decision-making process, which is highly important for adolescents and young adults to unfold in society.

3 Method

To examine the demand for this mobile application in society and how beneficial it can be, a quantitative descriptive study was conducted aimed at exploring the relevance of creating a mobile application to improve decision making in adolescents. To collect data on social perception, a questionnaire was applied to 153 adolescents and young

Table 1 Questionnaire frequencies

Item	Yes	Maybe	No	Didn't answer
1	103 (67.3%)	39 (25.5%)	9 (5.9%)	2 (1.3%)
2	74 (48.4%)	61 (39.9%)	12 (7.8%)	6 (3.9%)
3	87 (56.9%)	48 (31.4%)	15 (9.8%)	3 (1.9%)
4	55 (35.9%)	59 (38.6%)	37 (24.2%)	2 (1.3%)
5	69 (45.1%)	60 (39.2%)	22 (14.4%)	2 (1.3%)

people between 16 and 22 years old. The results show that this population finds useful and necessary a technological application to improve the decision-making process in the daily life of an adolescent. The results in detail are shown in the next section.

4 Results

The questionnaire included five items that allowed us to find out about the relevance of developing a mobile application aimed at improving decision making in adolescents. The proposed items were: (1) Do you think it necessary to develop technological devices to improve the neuropsychological functions of humans? (2) Do you consider that developing a mobile application for neuropsychological stimulation of decision making can help adolescents in their daily lives? (3) Do you consider that decision making can be improved if the subject has a cell phone application that provides information related to everyday situations and how to resolve them? (4) Do you think a teenager's decision making can be improved by using a technological device? (5) Do you consider it relevant to develop a mobile application so that a subject can improve their decision making? Table 1 and Figs. 1 and 2 show the results obtained from this questionnaire.

5 Conclusions

This research analyzed the existing demand for the creation of a mobile application (video game) that improves the decision-making process in the adolescent and young adult population. First, the prototype of the application was presented. Second, to explore the demand analysis a short questionnaire was answered by 153 people between 16 and 23 years old. Finally, the results were encouraging, as the perception indicates that the technological proposal may be needed and could bring benefits to the neuropsychological process in question.

It is important to mention that technology can be used for these purposes, resulting in an improvement in the quality of life and in determining neuropsychological

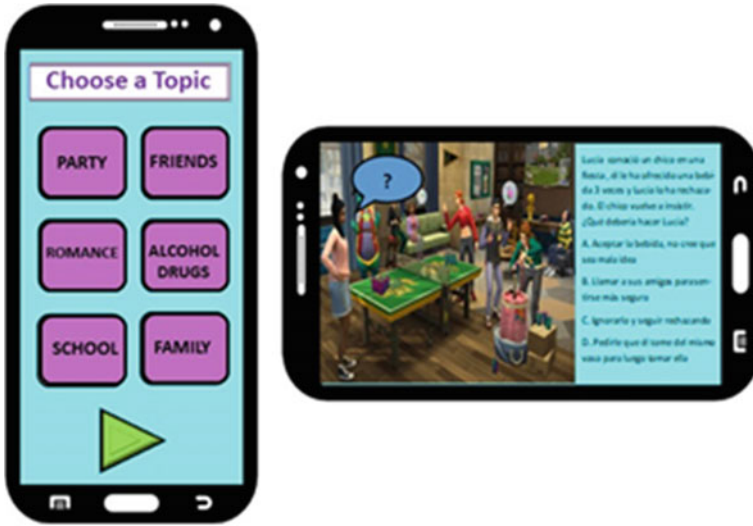


Fig. 1 Prototype of the mobile application to stimulate the decision-making process

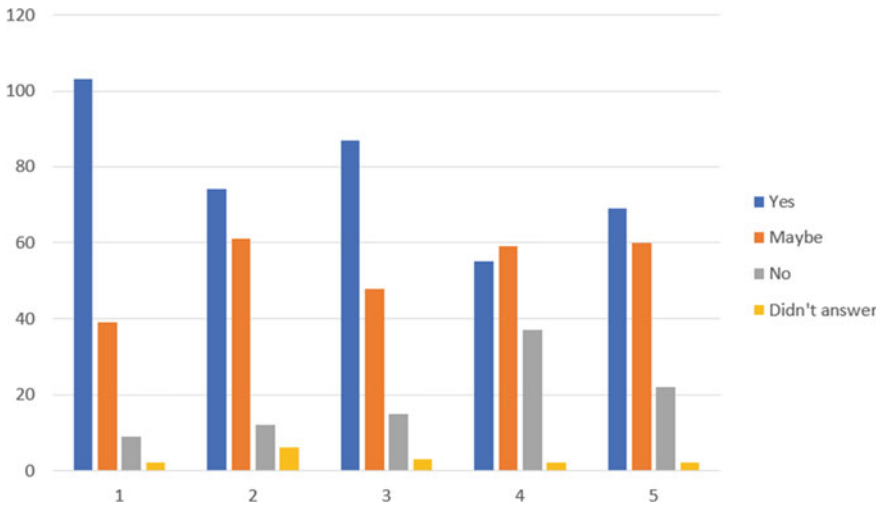


Fig. 2 Participants' responses diagram

processes such as decision making. Similarly, it is corroborated that a mobile application is of interest to the young population and is a more accessible way to train that process. It is also relevant to check whether the population feels that the objective of the proposal will be beneficial and target the technological research and development to the perceived social needs.

Future research is directed at the development of the conceptual proposal and the application of the mobile video game in the target population. This can provide ground for future research that analyzes the incidence and beneficial effects of the developed application. Likewise, technological intervention in other neuropsychological processes equal to or more important than decision making should not be set aside.

Acknowledgements Poner el proyecto que financia este trabajo.

References

1. Krauskopf, D.: El desarrollo psicológico en la adolescencia: las transformaciones en una época de cambios. *Rev Adolesc. Salud.* **1**(2) (1999)
2. Tamayo-Lopera, D., Merchán-Morales, V., Hernández Calle, J., Ramírez Brand, S., Gallo Restrepo, N.: Nivel de desarrollo de las funciones ejecutivas en estudiantes adolescentes de los colegios públicos de Envigado-Colombia. *Rev. CES Psico.* **11**(2), 21–36 (2018)
3. De Caro, D.M.: El estudio del cerebro adolescente: contribuciones para la psicología del desarrollo. V Congreso Internacional de Investigación y Práctica Profesional en Psicología XX. Jornadas de Investigación Noveno Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología, Universidad de Buenos Aires (2013)
4. Silva-Barragán, M., Ramos-Galarza, C.: Modelos de organización cerebral: Un recorrido neuropsicológico. *Rev Ec Neurol.* **29**(3). (2020)
5. Bohlin, M.C., Erlandsson, S.I.: Risk behaviour and noise exposure among adolescents. *Noise Health* **9**, 55–63 (2007)
6. Ramos-Galarza, C., Silva-Barragán, M., Condor-Herrera, O., Bolaños-Pasquel, M., Ramos, V., Cruz-Cárdenas, J.: Technological Resources for Neuropsychological Rehabilitation. *IHSI 2021, AISC 1322*, 181–186 (2021)
7. Aldana-González, G., García-Gómez, L., Jacobo-Mata, A.: Las Tecnologías de la Información y Comunicación (TIC) como alternativa para la estimulación de los procesos cognitivos en la vejez. *Rev. Inv. Ed.* **14**, 153–166 (2012)
8. Salazar-Mera, J., Gallegos-Nogales, L.: Aplicaciones móviles y su influencia en la flexibilidad cognitiva. Master's Thesis. Repositorio Digital Universidad Técnica de Ambato (2018)
9. Schwartz, D.M.: Breaking through barriers: using technology to address executive function weaknesses and improve student achievement. *Appl. Neuropsych. Child* **3**(3), 173–181 (2014)
10. Ramos-Galarza, C., Acosta-Rodas, M., Sanchez-Gordon, S., Calle-Jimenez, T.: Mobile Technological Apps to Improve Frontal Lobe Functioning. *AHFE 2020, AISC 1201*, 88–93 (2021)
11. Templeton, J., Poellabauer, C., Schneider, S.: Enhancement of neurocognitive assessments using smartphone capabilities: systematic review. *JMIR Mhealth Unhealth* **8**(6), e15517 (2020)
12. Bolaños-Pasquel, M., Silva-Barragán, M., Acosta-Rodas, P., Córdor-Herrera, O., Cruz-Cárdenas, J., Ramos-Galarza, C.: Technological Resources to Stimulate Multiple Intelligences: Verbal-Linguistic and Logical Mathematical. *IHSI 2021, AISC 1322*, 159–164 (2021)
13. Rossignoli-Palomeque, T.: Impacto del entrenamiento cognitivo NEXXO sobre la atención y funciones ejecutivas en edad escolar. Universidad Complutense de Madrid, Tesis doctoral (2020)

Neuromarketing: Current Applications in Favor of Consumerism



María Judith López and Carlos Ramos-Galarza

Abstract Neuromarketing is an innovative branch of neuroscience that seeks to provide for the consumer's wishes through the understanding of its brain structure, which will allow to comprehend and develop products or publicity that can be more attractive, new and attached to what has been created. Neuromarketing is integrated in this new digital age in an extraordinary way due to it having more opportunities to be applied. This science can be observed in each ad, which makes the individual feel the need to spend money on a product. Big companies have benefited from this almost-magic formula that allows the emotion and reasoning behind each purchase. Thanks to neuromarketing, each product could be attractive enough in the neuro-sensory aspect for it to be consumed, science being the only needed thing. This investigation provides the knowledge on the reason why it is one of the most researched sciences by the producer nowadays.

Keywords Neuromarketing · Neuropsychology · Marketing · Publicity · Consumer

1 Introduction

The evolution of sciences has started many knowledge and development fields, which in a group or individuals, offer the possibility of new perspectives toward the world and its reality. Searching for what the client wants and satisfying this has been a generational constant; hence, the market never ends and we, as human beings, can continue innovating in all the edges of science [1]. After this, the job of investigating consumer needs was born, and marketing became the science of observing. Now, the place where the link between a social science such as marketing and a health

M. J. López · C. Ramos-Galarza (✉)

Facultad de Psicología, Pontificia Universidad Católica del Ecuador. Av. 12 de Octubre y Roca, Quito, Ecuador
e-mail: caramos@puce.edu.ec

C. Ramos-Galarza

Centro de Investigación en Mecatrónica y Sistemas Interactivos MIST/Carrera de Psicología, Universidad Tecnológica Indoamérica. Av. Machala y Sabanilla, Quito, Ecuador

science such as psychology and neuropsychology started is what is being asked, and the answer stands in the central nervous system (CNS) and the peripheral nervous system (PNS), which are the fundamental pillars of neuromarketing [1].

The CNS is formed by the brain and the spinal cord, while the PNS is formed by the sensory receptors and the ganglia out of the CNS. This union of structures allows us as human beings to move the information we have obtained through our sensory organs, all the way from the skin, joints and muscles through the nerves to recognize and identify what we perceive and at the same time, the received information will have an emotional tint that will allow us to know if we like what was presented [1]. The brain, through its complex structures, can create this neuro-sensory path that transfers information, that is how a product's consumer processes his requirements and the company, businessman or producer, presents what is left of what needs to be bought and/or consumed.

Other fundamental structures of our brain that help marketing to propose sale standards are the brain hemispheres, the same that are divided in: occipital lobe (O.L), temporal (T.L), parietal (P.L) and frontal (F.L). The occipital lobe is the visual processing area; the temporal lobe is the sound, speech comprehension and memory processing one; the parietal lobe is the movement, orientation, calculus processing and certain recognizing area, and lastly the frontal lobe is the area of thought, concepts, planning and feeling control area [1]. Through the understanding of these structures, marketing makes the information processing through these lobes the clearest, most precise, direct, simple and attractive, for the consumer to be "hooked" to what has been sold or motivated to consume.

For marketing, the visual sense is one of the most significant senses and structures of study, because the recognizing and memory of the publicity image will be a decisive aspect on the product sales, the famous science of colors was born from it, theories such as Costanzo's [2] (see Chart 1) have existed, and they have allowed marketers to find the best way to sell their product that is at the same time attractive for the human brain [2]. On the other hand, Burkitt [3] cites the studies made by the multinational PepsiCo Inc. of Frito-Lay as an example through its publicity image, commercialized almost all over the world, in which they explain the way in which external stimuli related to the wrapping colors can influence the brain areas associated to guilt feelings in the consumers, added to the image of a natural or healthy ingredient in its package, decreasing the reluctance to consume this type of product in comparison to brilliant or stimulating colored packaging and little "natural" images of French fries [3] (Table 1).

The organs of the remaining senses such as hearing, taste, smell and touch are also an important part of the marketer's knowledge to analyze the efficiency of the product and thus be able to put it in the market. For marketing, the different qualities of sound such as tone, rhythm, intensity, volume, among others, represent vital elements at the moment of releasing a publicitary product because they will define the brand's positioning in the market [4]. In the field of publicity, the auditory stimuli that people perceive will be determinant as soon as feeling is produced in the hearer, allowing him to imagine and perceive more sensations, which invite him to check and most likely consume the product [4]. In terms of taste, it happens mostly through the taste

Table 1 Color theory

Color	Association	Product
Red	Movement, action, energy, activity, romance, warmth	Restaurants, children’s parks, entertainment centers
Blue	Tranquility, peace, coldness, relaxation	Dietary products, hospitals
Black	Status, elegance, power	Luxurious articles, high fashion, liqueur
White	Purity, innocence, cleanliness	Hygiene products and dairy products

buds in the mouth, nevertheless, smell is the contributing one for the system to work. Unfortunately these two systems haven’t been studied enough because they depend on a subjective component of people, what can be recognized are fragrances or flavors that can generally be liked by the population; however, this element motivates us to make further studies in the field of neuromarketing.

1.1 The Emotional Brain

It is common to hear, thanks to literature and poetry, how feelings and emotions are usually born in the heart, but as we could see before, the brain is the one that creates them and at the same time, has an effect on the remaining human being’s organs, depending on the way they’re perceived. There is a long neural lane that connects the thalamus which is the one that processes the sensory and motor information, and projects it to the brain cortex, with the cortex or thinking brain [5]. On the other hand, there is a lane that communicates the thalamus directly to the amygdala, whose principal role is the emotion processing and register. Therefore, there are two lanes, the first called the slow lane, and the second called the fast lane (Fig. 1). Now, what do these neural lanes have to do with neuromarketing? As proven, decision-making is strongly determined by the emotional system, it is clear to understand that successful companies in terms of publicizing the market are the ones that achieve the emotional impulse buy trigger and shortly after, fidelity to the brand when they continue consuming it impulsively [5].

Even though the human brain has separate structures to process the emotional and the rational components, as we just observed, both systems communicate and have an influence in the human behavior, therefore, even when there is a rational and judgmental opinion on what we consume, we almost always go to these two lanes, it has been observed that for the majority of the population, the emotional lane is usually the determinant one in the final choice [5]. For example, if someone is recommended a product that promises to change his life, and the recommender is much-loved, he himself will attempt to observe or analyze the product, if the person thinks it is too expensive, the emotional component of the recommendation and what is present in his life will be decisive in the final purchase.

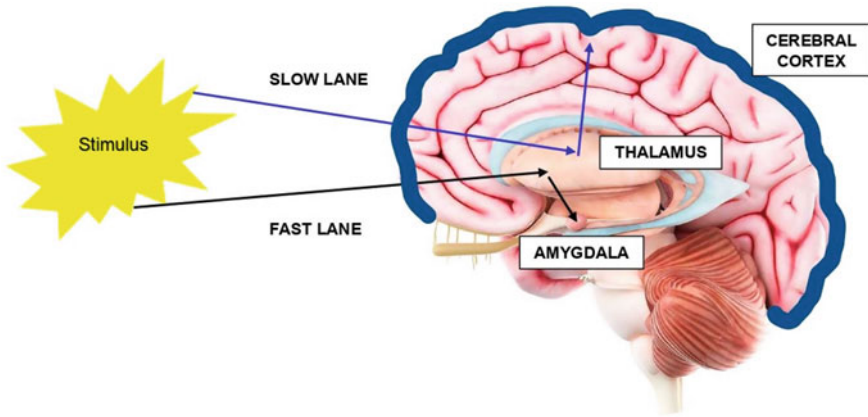


Fig. 1 Information processing lanes

1.2 Application of Neuromarketing

Next, we will mention the incidence of neuromarketing and its effectiveness, since at the time of analyzing the market, businessmen use different methodology and techniques to find the one that generates more income compared to the competition. It has been seen that the combination of neuropsychology, psychology and marketing has been one of the most successful ones.

1.2.1 Advertisement

Neuromarketing has been used to measure the effectiveness of publicity, with still images and videos played on television. There are many companies that have worked on the assessment on how to make your product sell more through neuroscience; for example, companies such as EmSense, helped Coca-Cola make their announcement in the Super Bowl more attractive in the neuro-sensory aspect [6]. On the other hand, Innerscope company, uses a one-sensor electrocardiogram, a breathing monitoring system and an accelerometer to measure the skin's galvanic response or electrodermal response; this measurement allows the variations of the skin's characteristics to be discovered thanks to certain types of sweat glands that indicate different emotions such as fear, wrath and others [6]. Knowing this allows the company to present, in a pilot test, different people and measure how much emotion is generated while watching the announcement, and this is added to the subjective response of each spectator, generating an almost-completely precise formula on the product's effect and the possible modifications that could be made for it to be striking, ornate and appealing to consume.

1.2.2 Multimedia Content

The neuropsychological data is widely used to evaluate emotional response; for example, 20th Century Fox used Innerscope to evaluate the trailers of their movies allowing the scenes that could cause strong emotion and make the viewer have a strong wish to watch and recommend the movie, to be chosen, generating a consumerism cycle. Microsoft used an electroencephalogram to study the possibility of recognizing simple tasks done by the users measuring the brain activity. This allowed them to analyze whether it is easy for new users to learn how to use the operative system that allows them to interact easily in the interface [6]. On the other hand, the use of social media is more of an escape from reality; it is true that what is consumed because of this influence comes from posts from friends or family, but now, it is seen as a new place to sell and consume products, publicity videos between songs and images are daily situations that could cause discomfort but are inviting consumerism because of the innovation they represent, where our brain applies the fast lane or the slow lane to decide on the product [6].

1.2.3 Video Games

The population has been modified, now, new generations are usually raised in a technological and innovative world, where videogames are the new method for many teenagers and adults to have fun, socialize and be entertained. We can see, more than ever, how the emotion in video games is advertised to attract new teenagers to consume them. Thanks to the technological innovation in the image quality and controls, publicity has become easier for marketers, making teenagers and adults feel a type of necessity and almost addiction to what they consume, which at the same time allows big companies to profit from this. Neuromarketing has had an alarming effect on videogames, and its effectiveness has been proven in the addictions created toward them, a new obsession to purchase the new technological novelty has been generated. Internet addiction occurs when the user stops doing all daily activities in order to play video games. This represents the success and the consequences of marketing on video games sales and distribution [7, 8].

2 Conclusions

Neuromarketing is part of the wide range of neurosciences. The study of the multidisciplinary brain, due to its various levels of study, from the molecular one, to the cognitive and conductual one in human beings, with this, the study of neurosciences is proven to be complex, but necessary to better understand human functioning. Thanks to the extensive current knowledge on the human brain, neuromarketing has been able to take its place in the world at a fast pace. This shows that with the correct formula combining both neuropsychology and marketing, we could discover the consuming

conduct of big populations. It is not a surprise that this is already happening, since now, neuromarketing has shown consumerism approaches that people do not need, although many companies make us believe they do, through the presentation of color, smell sound and the represented experience. Many consumers do not buy the product because of what it signifies materialistically, but because of what it represents for society, whether that be status, “good taste” or being in the latest trends.

References

1. Virginia, M., Del Fatima León, M., Mora, C.: Neuromarketing: Know Customers Through Their Perceptions. *s Tec Empresarial*, pp. 9–19 (2010)
2. Constanzo, L.: *Physiology*. Mc.Graw Hill Interamericana, Madrid (2006)
3. Burkitt, L.: Battle for the brain. *Forbes* **184** (2005)
4. Robert, K.: *The Future Beyond Brands*. Lovemarks, Barcelona (2005)
5. Braidot, N.: *Neuromarketing*. Ediciones Gestion (2005)
6. Huespe, A., Rotellini, M., Squarzon, M., Tartarelli, E.: Neuromarketing, setting a path. *Hologrammatic*, pp. 239–261 (2014)
7. Echeburua, E.: Addiction to new technologies and to online social networking in young people: a new challenge. *Addictions*, pp. 91–95 (2010)
8. Lopez, M., Moscoso-Salazar, J., Ramos-Galarza, C.: *Technological Innovations in the Treatment of Addictions* (2021)

Expectations versus Perceptions: The Role of Consumers' Senses in Composing of USP for PDO Wines



Jiří Zelený , Karolina Macháčková , Jan Hán , and Michal Lošťák 

Abstract The unique selling proposition (USP) as a major distinctive feature has been studied in various marketing researches. Our study uses sensory evaluation, a method introduced in many studies employing neuromarketing, to uncommonly examine the USP. As an example, Czech wines with the PDO designation called VOC were selected. The sensory evaluation was complemented by price analysis and qualitative research. First, the price analysis that included 805 Czech wines showed that the VOC wines are less common on the Czech market and have higher prices than substitutes without the VOC. Second, the blind sensory evaluation done by 103 wine consumers underlined the relative perception of wines and their USPs. The VOC was not perceived as a decisive attribute for consumers choosing from the VOCs and non-VOCs. The absence of USPs for VOCs from the consumers' point of view shows the relative perception of USPs by various agents (producers, consumers). Third, the semi-standardized interviews with 15 volunteers from sensory evaluation explored consumers' expectations regarding VOC wines. The experienced consumers matched the VOCs with distinctive terroir and taste, i.e., possession of USP, which opposed their actual sensory perception. We believe that the consumers' sensory evaluation should be an elementary part of marketing research, including creating food product's USP.

Keywords Czech Republic · Qualitative research · Sensory evaluation · Terroir · VOC wines · Willingness to pay

J. Zelený (✉) · K. Macháčková · J. Hán
Institute of Hospitality Management, Svédnická 506, 181 00 Prague, Czech Republic
e-mail: zeleny@vsh.cz

M. Lošťák
Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýcká
129, 165 00 Prague, Czech Republic

1 Introduction

In numerous fields, long-term and contemporary research attention have been paid to unique selling proposition (USP), a frequently used marketing term [1–3]. The USP stands for a clear benefit that should be provided to a customer when he/she decides to purchase the product of a certain company instead of purchasing a substitute from competition [4]. Therefore, the USP is a distinctive feature and an element that fundamentally influences the company's communication with its customers through advertisement activities [5]. Our study addresses “distinctiveness” as an essential attribute that defines the USP [1].

The vast majority of previously published studies on USP used varied marketing or business models, but our paper shows the less common use of sensory analysis in marketing. Thus, our research aims to reveal the necessity of consumers' sensory evaluation usage in creating successful USP. An example of Czech wines with protected designation of origin (PDO) called VOC (from the Czech “Vína Originální Certifikace” meaning Wines of Original Certification) was used to fulfill the aim. The research uses a combination of quantitative and qualitative methods to highlight quality relativism that should instruct the producers when deciding how their USP should be constructed.

2 Literature Review

Concerning the focus of our study, the definition of USP applies to food and beverages as well and is of increasing importance, especially for food market segments with growing numbers of products. Such uniqueness may be provided, for example, via organic labels affecting consumers' perception of health food products, as shown by Bauer et al. [6] or Massey et al. [7]. For published articles for which the object of study has been wine and viticulture, the USP (uniqueness/distinctiveness, respectively) is connected not just with the term “brand” [8] but also with words as “authenticity” [9] or “terroir” [10].

The conception of terroir vigorously represents the USP in the wine and viticulture field since some authors consider terroir a “distinguishing” factor. In particular, terroir is a collection of various environmental conditions of the origin of grapes, which should subsequently be reflected in the unique sensory characteristics of wines [11, 12]. A wine that appropriately reflects terroir should, therefore, have a well-established USP. However, formalizing the concept of terroir in an understandable way for consumers is not easy. In the European Union, the best performing classifications that formalize the terroir use the abbreviations of protected designation of origin (PDO) and protected geographical indication (PGI), concepts guaranteeing the geographical (regional) origin of an agricultural product, its uniqueness, and tradition [13, 14]. The previous research using survey data has shown that decision-making according to both PGI and PDO is essential for Italian consumers [14, 15].

While examining the influence of terroir or PGI/PDO on consumer attitudes, the marketing survey approach has been widely used in the studies mentioned above. However, examining the USP of wines with PGI/PDO using sensory analysis is unusual and pioneering even in such wine countries as Italy [15]. Sensory analysis using the consumers' senses is often employed in neuromarketing studies as it can adequately uncover consumers' preference patterns that may not appear in questionnaires [16]. The previous research has shown that the smell of various products [17], such as wines [16], is essential for consumers. On the other hand, revealing of hidden patterns during sensory evaluation, e.g., when choosing between two PGIs, can show that the consumers cannot find any differences. Thus, consumers' attitudes revealed in the questionnaires can be significantly reduced when sensory evaluation takes place [15].

Our study uses a sensory evaluation approach employing consumers but does not compare the two PGIs as the previous studies [15]. Instead, it compares even more different categories—wines with the PDO designation and wines without PDO or PGI. Specifically, the research is performed on wines with the designation VOC as the Czech equivalent of PDO wines. The VOC system is a newer variant compared to the traditional Germanic wine labeling system in the Czech Republic, and it has been used since 2009. According to the Wine Fund of the Czech Republic, VOC wines should bear typical regional features, i.e., the wines should reflect the terroir [18]. The winemakers themselves approve each wine with the VOC designation during a blind sensory evaluation. Therefore, each approved VOC wine should correspond to the winemakers' conception of the ideal regional specificity. At present, there are 13 regions in the Czech Republic bearing the VOC designation.

3 Research Methodology

The research in this study includes two PDOs, namely VOC Znojmo and VOC Mikulov. The first PDO is the oldest one and has been creating the USP for the longest time. VOC Mikulov is the most extensive PDO as it belongs among the three PDOs associating the most winemakers, and at the same time, it belongs to the three PDOs with the largest number of permitted grape varieties for the production of VOC wines. The example of the Czech Republic was chosen as the country where both the PDO wine system and traditional Germanic wine labeling exist. In this respect, the Czech Republic differs from other countries and is in a unique position. Another suitable example for a similar examination would be Austria, where both systems are present too. At the same time, the Czech Republic was selected as the country for which data were available, i.e., a demonstrative example underpinning examined theoretical conceptions.

In the first phase of the research, the price level of wines sold by winemakers belonging to the two examined VOCs was examined. Specifically, 805 wines, their price, and disposition of PDO were recorded in the e-shops of winemakers. The price comparison then concerned wines with the VOC designation and wines without the

VOC designation (produced from varieties permitted for VOCs). The Mann–Whitney U test ($\alpha = 5\%$) was used for comparison due to the smaller size of the groups and due to the non-parametricity of the data tested by the Kolmogorov–Smirnov test ($\alpha = 5\%$). Before the statistical testing took place, the outliers were removed from the groups according to Tukey’s method using the interquartile range (IQR) [19]. Values exceeding the IQR by 2.2 times were removed as values exceeding the IQR by 1.5 times were considered insufficient, according to Hoaglin and Iglewicz [20]. Such testing is feasible despite the different sizes of the groups, as demonstrated by Zimmerman [21], who tested unequally large groups in a ratio of 1:4.

The second phase consisted of a sensory evaluation of wines. Specifically, a comparison of wines with and without VOC labeling was performed to assess the perception of the USP from consumers’ point of view. A total of 103 wine consumers over 18 years old who also buy wine regularly took part. Male consumers (46.53%) and female consumers (53.47%) were present. The youngest consumer was 18 years old; the oldest was 71 years old. The mean age of the consumers was 36.62 ± 14.63 . Consumers had both a secondary education (44.90%) and university education (55.10%), and both students (43.43%) and employed participants (56.57%) took part. In terms of average monthly income in the Czech Republic, there were consumers with above-average (38.90%), average (35.56%), and below-average (25.56%) income.

The relatively small number of participants in sensory evaluation (e.g., in comparison with a standard questionnaire survey) is due to a more complex logistical provision of tasting. However, the advantage of sensory evaluation is the revelation of the actual preference of consumers, which is not only based on the verbal or visual presentation of the product, but it is also based on assessing product appearance, aroma, and taste (or touch via *nervus trigeminus*). Even a relatively small number of participants is considered sufficient during the sensory evaluation by the professional literature, as stated, for example, by Martišová et al. [22], who used ten expert evaluators. For untrained consumers, the required number is higher, but the rule of the central limit theorem has been applied [23, 24], i.e., at least 30 evaluators, which is in line with Blanchera et al. [25]. The number of evaluators in this study is similar to those performed with 109 [26], resp. 131 evaluators [27].

Four wines were selected for evaluation. The characteristics of these wines are given in Table 1. Even in this case, compared to some purely sensory studies, the number of samples can be considered smaller. Nonetheless, mixed methods are used in our study, and sensory evaluation is extended by the economic concept of willingness to pay (WTP) and qualitative research. Previously performed economic-sensory studies also included only four evaluated samples [26]. There were two separately evaluated pairs in our study, with one of the pairs coming from the Znojmo region and the other from the Mikulov region. The main difference within each evaluated pair was the classification of wines—one sample was with a VOC designation, and the other without VOC designation. The purpose was not to compare all four wines (particular VOCs vary greatly) but to explore the preferential patterns for each pair of wines, with the wines in the pair being considered substitutes with the same winery,

Table 1 Characteristics of evaluated wines samples

Sample No. (CODE)	1 (LED)	2 (ZNO)	3 (STR)	4 (SON)
Region	Znojemsko	Znojemsko	Mikulov	Mikulov
Village	Krhovice	Krhovice	Popice	Popice
Vineyard	Skalka	Skalka	unknown	Sonberk
Winery	Znovín	Znovín	Sonberk	Sonberk
Variety	Gr. Veltliner	Gr. Veltliner	Riesling	Riesling
Vintage	2018	2018	2017	2017
Sugar	7.3 g/l	6.9 g/l	dry	4.6 g/l
% alc	12.5	12.5	13.0	12.5
Acids	6.9 g/l	6.7 g/l	Unknown	8.0 g/l
VOC	Absent	Present	Absent	Present

Pictures of evaluated wines



variety, and vintage. The selection of specific wines was due to the suppliers' possibilities. Only wineries that offered wines with the same analytical properties but different in the presence/absence of the PDO were selected.

The evaluation of the wines consisted of two phases—Phase I was a visual evaluation of the appearance of the bottles based on the projected photographs—initially, the first pair of wines and then the second pair of wines. Consumers indicated their WTP (in CZK) for each wine and were reminded that the wines in the pair could be considered substitutes. If consumers indicated different WTPs for wines, they were asked to state the reason for the higher/lower WTP.

Phase II consisted of sensory assessment, i.e., evaluating four wines from Phase I, but poured into glasses so that consumers could use the sense of sight, smell, and taste to assess intrinsic attributes. In order not to influence consumers during the sensory evaluation by the judgment based on the presented photographs of wines, the wines were poured before the arrival of the consumers. Consumers were warned that the tasted wines might be different from those presented in Phase I. Sensory

evaluation also compared two pairs of wines (two pairs of separate glasses). Even in this case, consumers were warned that each pair consists of wines that can be considered substitutes. However, only the order of the evaluated wines in pairs was reversed compared to Phase I. In addition, the sensory evaluation included a standard 9-point numerical hedonic scale of acceptability in the form of a semantic differential, where consumers stated their attitudes toward the wines being tasted, from completely unacceptable/worst to completely acceptable/best (according to ISO 6658) [28]. Indication of the WTP for each wine followed with reasons for possible different WTP in pair of wines. The test for independent groups, i.e., the Mann–Whitney U test, was used to detect changes during Phase II compared to Phase I. The Wilcoxon test for dependent groups was used for testing wines with each other only within one of the phases. All tests were performed at a significance level of $\alpha = 5\%$.

In the third phase of the research, semi-standardized interviews were conducted with 15 consumers who participated in sensory evaluation and subsequently volunteered for the qualitative research. During the interviews, the attitudes of consumers (especially their expectations) for wines with the VOC label compared to wines without the VOC label were explored. Disclosure of individual wine samples was sent to participants by email so that participants who attended the interviews were not influenced during the interviews. The interview participants were briefly introduced to the principle of VOC labeling if some of them were not familiar with it.

4 Results and Discussion

Based on the results of descriptive statistics, it can be stated that the wineries, despite their membership in the VOC designation, offer the majority of wines classified outside the VOC designation (70.06%). In this case, these are wines for which grape varieties not authorized to produce VOC wines have been used. A smaller part of the offer (24.35%) is formed by wines produced from varieties authorized for VOC wines but not bearing the VOC label. The wines with the VOC label themselves form only a negligible part of the wineries' offer (5.59%), and it is possible to assume their uniqueness and disposition of the USP compared with other products. Therefore, in terms of the breadth of the product range, wines with the VOC designation cannot be considered decisive. Subsequent testing presents price differences between VOC and non-VOC wines, which would potentially be suitable for the production of VOC wines, as only such a comparison of substitutes is relevant. The results are shown in Table 2, where VOC-labeled wines are significantly more expensive ($U = 3249.00$; $p < 0.01$) than non-VOC-labeled wines. Overall, it can be stated that VOC wines are less represented on the Czech market, and they are significantly more expensive than production without VOC labeling, even if it is made from the same grape varieties. Therefore, we consider VOC production to be more premium in price and a smaller part of the wineries' portfolio. Thus, VOC wines might be more suitable for possessing USP.

Table 2 Portfolio composition and prices of wines for investigated wineries

Classification type	VOC wines	No VOC; varieties permitted for VOC production	No VOC; varieties forbidden for VOC production
Number of wines	45	196	564
Rel. count (%)	5.59	24.35	70.06
Median (CZK)	250.00	200.00	
Mean \pm sd (CZK)	239.64 \pm 59.50	212.37 \pm 73.16	
Mann–Whitney U	$U = 3249.00; Z = -2.63; p < 0.01$		

Note: Presented prices are in Czech koruna (CZK), where currency conversion for July 2021 was 1 USD = 21.70 CZK

Further results (see Table 3) show how the customers-evaluators themselves react to the premium products of the wineries that are represented by VOC wines. Statistical analysis showed that consumers' WTP always correlated significantly with their proclaimed hedonic acceptability of wines during the sensory evaluation, even at $\alpha = 1\%$. Therefore, the results section contains only the presentation of WTP results for the visualized/tasted wines. While for wines from the Mikulov region, no statistically significant differences were recorded in the evaluation of the WTP for Phase I and II ($T = 629.00; p = 0.17$ and $T = 691.50; p = 0.28$), significant differences were recorded in the evaluation of wines from the Znojmo region, moreover different in individual phases. During the visual evaluation of wines, consumers rated significantly better ($T = 387.50; p < 0.01$) the ZNO wine (219.49 ± 39.70 CZK) than the LED wine (205.64 ± 36.13 CZK). On the contrary, after wine tasting, the consumers significantly better rated ($T = 423.00; p = 0.01$) the LED wine ($T = 222.05 \pm 43.05$ CZK) opposed to ZNO wine (212.86 ± 37.09 CZK). Compared to the previous studies [14, 15], our results indicate that the presence of PGI does not engage the consumers to choose such a wine in all cases during the visual evaluation of the bottles. The same situation appears when the wines are tasted, as consumers can prefer wines without the PDO designation. Therefore, the perceptual gap between wines depends on a particular sample.

Table 3 Results of consumers' sensory evaluation

Statistic/wine	Phase I (mean \pm sd)	Phase II (mean \pm sd)	Mann–Whitney U (Phase I + II)
LED	205.64 \pm 36.13	222.05 \pm 43.05	$U = 73.50; p = 0.49$
ZNO	219.49 \pm 39.70	212.86 \pm 37.09	$U = 135.50; p = 0.60$
Wilcoxon T	$T = 387.50; p < 0.01$	$T = 423.00; p = 0.01$	
SON	223.68 \pm 41.60	201.03 \pm 34.51	$U = 159.50; p = 0.09$
STR	228.13 \pm 44.13	204,18 \pm 35,25	$U = 157.00; p = 0.73$
Wilcoxon T	$T = 629.00; p = 0.17$	$T = 691.50; p = 0.28$	

Consumers choosing a higher WTP for the ZNO wine did so in 54.76%, mainly due to the displeasure of the LED wine logo. This result indicates more unpopularity of the LED wine logo than the more significant popularity of the VOC logo. It is also in line with the results of the WTP comparison in Phase I for wines from Mikulov, where the VOC logo was not the reason for the SON wine's higher popularity (although the logo of the SON wine has a slightly different appearance compared to the ZNO wine). Reasons for higher WTP in Phase II for the LED wine were limited only to a general reason of "taste" (31.91%), although regular wine consumers participated in the sensory evaluation. Similarly, consumers more often choosing a higher WTP for the ZNO wine stated "taste" as their main reason (37.93%). The second most frequently cited reason for higher WTP was lower acidity or higher sweetness (23.40% for the LED wine and 20.69% for the ZNO wine). However, Table 1 shows that wines differed in the residual sugar content by 0.4 g/l and in the acid content by 0.2 g/l. This result testifies to a subjective perception of individual components of wines (which are objectively comparable for wines in terms of content), even for regular wine consumers.

When examining each wine separately, it cannot be said that a change in the evaluation method, i.e., the transition from Phase I to Phase II, would mean a change in the WTP. Thus, our study shows different knowledge than papers underlining discrepancies between questionnaires and sensory evaluations [16]. In other words, there was no significant change in the willingness to spend money after the wine tasting ($p = 0.09$ to $p = 0.73$). Therefore, we believe that the possible perception of the wine label uniqueness is in accordance with the perceived uniqueness of wines' tastes. The consumers' point of view approves that the winemakers were most probably able to balance the USP of bottle appearance with the taste profile of the wine. Change of preferred wine in Phase I and Phase II in the case of the Znojmo region indicates a change that does not occur per se but is always in comparison with other wines. It is necessary to emphasize the relativity of the evaluation, i.e., an attitude is established toward a pair of wines, not to a single wine. It cannot be said that after Phase II, consumers would be willing to pay significantly more for the LED wine, just as it cannot be said that after tasting, consumers would be willing to pay significantly less for ZNO wine. The change is to be understood as such, which means prioritizing the label of the ZNO wine over the LED wine, and further prioritizing the appearance, smell, and taste of the LED wine over the ZNO wine. Therefore, the USP of wine is most probably perceived relatively, i.e., with regard to a USP of other wine.

According to quantitative results, the more expensive and less common Czech VOC as an example of PDO does not seem to possess a well-elaborated USP, at least from the consumers' point of view. Surprisingly, in one of the cases, the consumers preferred significantly more the wine without the VOC designation over the substitute with the VOC designation during the tasting. In the second case, the quality of the substitutes (VOC and non-VOC) was perceived indifferently. It applies to a visual evaluation of the bottle, the appearance, aroma, and taste of VOC wines.

The results of the third phase of the research, which included interviews with 15 consumers, showed two main ideal types of consumers. The first group included less

aware participants consuming wine. Understandably, these consumers did not know the VOC and approached it somewhat reluctantly, even after a short introduction. In particular, they considered VOC wines to be those that would not be rated well by them as less experienced consumers. At the same time, they could not assess whether such product is different from other premium production without a VOC label. Thus, they could not assess the presence of the USP well, and their expectations regarding VOC were completely absent. Examples of selective transcription are given below:

I probably would not recognize the VOC during the tasting, so if it was expensive, I would not buy it - I am not a gourmet, so I do not care (C2).

... If it (authors' note: he talks about VOC) will be next to French wine in a shop, I will buy the foreign one. It will probably be better, but I have never tasted it, so I do not know what the advantage is for me (C7).

However, the second group of more knowledgeable consumers who already had experience with VOC showed opposite tendencies to the first group of consumers. They considered VOCs to be something with clear expectations of uniqueness and differentiation from standard production. Their view corresponds to a conception seen through the prism of marketing as a well-elaborated USP of product. Further examples of selective transcription are given below:

I know VOC well. I associate it with higher quality compared to other Czech wines. When I buy the wine, I expect that there will be an expression of terroir, which the competition will not be able to offer (C8).

When it has a VOC, it will not be standardized production like some large wineries. When it has a VOC, I associate it with something that has a different taste (C11).

Based on the statements provided by experienced consumer C8, the terroir is being related to an advantage that competition's everyday wines cannot offer. Therefore, it can be agreed that distinctive terroir is necessary for successful USP [4, 10]. The message provided by the interviewed experienced Czech consumers is evident. They expect VOC to provide USP. Nonetheless, the sensory evaluation results showed that the actual perception of the consumers is that VOC wines are in terms of a quality indistinguishable or worse than substitutes without the VOC designation. To control these results, we ran the statistical tests once more for the group of more experienced evaluators, and there were no significant differences compared to the results in Table 3. In conclusion, the winemakers should reconsider the USP of VOC wines. Marketing research and sensory evaluation with consumers can be helpful in the search for USP of Czech PDO wines.

5 Final Remarks

Based on the presented results, sensory evaluation by consumers can be considered an elementary part of marketing research, including the creation of wine USP. However, the results are most likely to be generalized to other than wine products.

The role of the senses (as approved by other studies [16, 17]) is crucial in creating the USP because it amplifies the differences between the objective properties of products (which may raise certain expectations) and their subjective perception. In addition, our study evokes other questions concerning, for example, how different is the perception of USP for individual groups, i.e., producers and consumers (in our study for winemakers and consumers). Therefore, marketing managers should not directly neglect the consumers' sensory evaluation of products, as managers' sensory perception may be subject to operational blindness.

Our research followed the other studies [26, 27] employing a limited number of regular sensory evaluators. Only a few evaluated wines can be considered as another limit of research, but it had been used in previously published studies as well [26]. Nonetheless, selected wines within each pair are ideal substitutes offered by wineries, and it should be noted that approved VOC wines are subjected to a blind sensory test by a winemakers' commission. Therefore, all VOC wines (including the wines rated in our study) represent an approved quality perception of the winemakers. In addition, the first selected winery in our study is among the ten largest wine producers in the Czech Republic. The second winery was repeatedly awarded as the winner of the Winery of the Year competition. Future research may be extended to more VOCs or examine different consumers' clusters and attitudes toward USPs of wines in the Czech Republic or other countries. From a theoretical perspective, our study shows that even today, relativism is still relevant in marketing science, and the crucial issue is the understanding of the differences in the attitudes of individual agents [29].

Statements and Acknowledgement The Ethics Committee of the Czech University of Life Sciences Prague has approved all research methods of this study and the attendance of all participating volunteers. The authors disclosed the financial support by The Ministry of Culture of the Czech Republic, project No. DG18P02OVV067 entitled Culinary Heritage of the Czech Lands: Memory, Presentation and Education.

References

1. Niu, Y., Wang, C.L.: Revised unique selling proposition: scale development, validation, and application. *J. Promot. Manage.* **22**(6), 874–896 (2016). <https://doi.org/10.1080/10496491.2016.1214209>
2. Tarman, H.A., Soleh, S.M., Ari, A., Rahmat, T.A.: Leveraging brand equity by applying brand communication and forming city branding based on unique selling proposition (A case of Crafts city). *Int. J. Bus. Adm. Stud.* **5**(2), 74–83 (2019)
3. Wang, Y., Huang, L., Li, J., Yang, Y.: The mechanism of tourism slogans on travel intention based on unique selling proposition (USP) theory. *J. Travel Tour. Mark.* **36**(4), 415–427 (2019). <https://doi.org/10.1080/10548408.2019.1568950>
4. Kippenberger, T.: Remember the USP? (Unique selling proposition). *Antidote* **5**(6), 6–8 (2020). <https://doi.org/10.1108/eum000000006861>
5. Dens, N., Pelsmacker, P.D.: How advertising strategy affects brand and USP recall for new brands and extensions. *Int. J. Advert.* **29**(2), 165–194 (2010). <https://doi.org/10.2501/s0265048710201117>

6. Bauer, H.H., Heinrich, D., Schäfer, D.B.: The effects of organic labels on global, local, and private brands. *J. Bus. Res.* **66**(8), 1035–1043 (2013). <https://doi.org/10.1016/j.jbusres.2011.12.028>
7. Massey, M., O’Cass, A., Otahal, P.: A meta-analytic study of the factors driving the purchase of organic food. *Appetite* **125**, 418–427 (2018). <https://doi.org/10.1016/j.appet.2018.02.029>
8. Pugh, M., Fletcher, R.: Green international wine marketing. *Australas. Mark. J.* **10**(3), 76–85 (2002). [https://doi.org/10.1016/s1441-3582\(02\)70159-6](https://doi.org/10.1016/s1441-3582(02)70159-6)
9. Griessbach, L., Ghvanidze, S., Lakotta, J.: Benefits convincing wine consumers—developing a unique selling proposition for Georgian wines based on focus group analysis. In: 9th Academy of Wine Business Research Conference, February 17–18, 2016, pp. 135–144. Adelaide (2016)
10. Leriche, C., Molinier, C., Caillé, S., Razungles, A., Symoneaux, R., Coulon-Leroy, C.: Development of a methodology to study typicity of PDO wines with professionals of the wine sector. *J. Sci. Food Agric.* **100**(10), 3866–3877 (2020). <https://doi.org/10.1002/jsfa.10428>
11. Matthew Rendleman, C., Hoemmen, G.A., Altman, I., Taylor, B., Moon, W., Smith, S.: Wine industry competitiveness: a survey of the Shawnee hills American viticultural area. *Wine Econ. Policy* **5**(1), 4–13 (2016). <https://doi.org/10.1016/j.wep.2016.03.002>
12. Souza Gonzaga, L., Capone, D.L., Bastian, S.E.P., Jeffery, D.W.: Defining wine typicity: sensory characterisation and consumer perspectives. *Aust. J. Grape Wine Res.* **27**(2), 246–256 (2020). <https://doi.org/10.1111/ajgw.12474>
13. Chandra, S., Chapman, J., Power, A., Roberts, J., Cozzolino, D.: Origin and regionality of wines—the role of molecular spectroscopy. *Food Anal. Methods* **10**(12), 3947–3955 (2017). <https://doi.org/10.1007/s12161-017-0968-1>
14. Di Vita, G., Caracciolo, F., Brun, F., D’Amico, M.: Picking out a wine: consumer motivation behind different quality wines choice. *Wine Econ. Policy* **8**(1), 16–27 (2019). <https://doi.org/10.1016/j.wep.2019.02.002>
15. Galletto, L., Caracciolo, F., Boatto, V., Barisan, L., Franceschi, D., Lillo, M.: Do consumers really recognise a distinct quality hierarchy amongst PDO sparkling wines? The answer from experimental auctions. *Br. Food J.* **123**(4), 1478–1493 (2020). <https://doi.org/10.1108/bfj-07-2020-0625>
16. Horska, E., Bercik, J., Krasnodebski, A., Matysik-Pejas, R., Bakayova, H.: Innovative approaches to examining consumer preferences when choosing wines. *Agric. Econ. (Zemědělská Ekonomika)* **62**(3), 124–133 (2016). <https://doi.org/10.17221/290/2015-agricecon>
17. Rosenlacher, P., Tichý, J.: The effect of sensory marketing from the perspective of neuromarketing. *GRANT J.* **9**(1), 77–80
18. Vinařský fond (Wine Fund): Co znamená označení VOC? (What VOC designation means?). Retrieved from <https://www.vinazmoravyvinazcech.cz/cs/encyklopedie/jak-se-vyznat-v-nasich-vinech/co-znamenava-oznaceni-voc-ds/co-znamenava-oznaceni-voc> (2021). Accessed 10 Sept 2021
19. Tukey, J.W.: *Exploratory Data Analysis*. Addison-Wesley Pub. Co., Reading, Massachusetts (1977)
20. Hoaglin, D.C., Iglewicz, B.: Fine-tuning some resistant rules for outlier labeling. *J. Am. Stat. Assoc.* **82**(400), 1147–1149 (1987). <https://doi.org/10.1080/01621459.1987.10478551>
21. Zimmerman, D.W.: Comparative power of student T test and Mann-Whitney U test for unequal sample sizes and variances. *J. Exp. Educ.* **55**(3), 171–174 (1987). <https://doi.org/10.1080/00220973.1987.10806451>
22. Martišová, P., Štefániková, J., Hynšt, M., Vietoris, V., Macková, P.: Preference mapping of Slovak cheese. *Potravinárstvo Slovak J. Food Sci.* **14**, 1211–1217 (2020). <https://doi.org/10.5219/1092>
23. Avery, P.J., Masters, G.A.: Advice on the design and analysis of sensory evaluation experiments using a case-study with cooked pork. *J. Roy. Stat. Soc. Ser. D (The Statistician)* **48**(3), 349–359 (1999). <https://doi.org/10.1111/1467-9884.00193>
24. Minoza-Gatchalian, M., de Leon, S.Y., Yano, T.: Quantified approach to sensory panelist selection. *Food Qual. Prefer.* **2**(4), 233–241 (1990). [https://doi.org/10.1016/0950-3293\(90\)90015-m](https://doi.org/10.1016/0950-3293(90)90015-m)

25. Blancher, G., Clavier, B., Egoroff, C., Duineveld, K., Parcon, J.: A method to investigate the stability of a sorting map. *Food Qual. Prefer.* **23**(1), 36–43 (2012). <https://doi.org/10.1016/j.foodqual.2011.06.010>
26. Tozer, P.R., Galinato, S.P., Ross, C.F., Miles, C.A., McCluskey, J.J.: Sensory analysis and willingness to pay for craft cider. *J. Wine Econ.* **10**(3), 314–328 (2015). <https://doi.org/10.1017/jwe.2015.30>
27. Teuber, R., Dolgoplova, I., Nordström, J.: Some like it organic, some like it purple and some like it ancient: consumer preferences and WTP for value-added attributes in whole grain bread. *Food Qual. Prefer.* **52**, 244–254 (2016). <https://doi.org/10.1016/j.foodqual.2016.05.002>
28. Bower, J. A.: *Statistical methods for food science: introductory procedures for the food practitioner*. Wiley-Blackwell, Ames, Iowa (2009).
29. Muncy, J.A., Fisk, R.P.: Cognitive relativism and the practice of marketing science. *J. Mark.* **51**(1), 20 (1987). <https://doi.org/10.2307/1251141>

Enterprise Information System for Digitalization and Organization of Innovation—A Case Study



Zornitsa Yordanova 

Abstract Organizational innovativeness and firm's abilities to innovate rely exclusively on team efforts, culture for innovation and senior management support, all these organized into a structured process. Organizational innovativeness has a variety of manifestations, elements, and sub-processes for cultivating and gathering individual and group abilities for the development of innovations within the organization. It represents a separate and parallel innovation process to the traditional innovation development process. The organization of this process and its digitalization as well as building long-term and continuous improvement through innovation mind-set is a challenging exercise for enterprises. In this study, we present a case study with the platform GAPideaz—an enterprise Web system for organizing and digitizing a simple innovation process encompassing all employees of an enterprise and focusing team efforts for the purpose of innovation development and problem solving. Major research and practical gaps in this organizational innovation effort are identified and addressed. The study contributed to the analyzed platform to precise its process and users target and to some of the users involved in the study to better organize the impact of the platform for organizing innovation across their enterprise.

Keywords Innovation management · Digital transformation · Enterprise information system · Organizational innovation

1 Introduction

Digitalization of enterprise innovation processes has been largely discussed since the beginning of 90s when the most complex business processes have started to be atomized. Among the pioneers in this field is Swanson [1] who analyzed information systems as a means for innovation and their crucial role for the organizing innovation circuit in an enterprise. Information systems have been proposed as a reference discipline for the development of other fields as well [2] since they

Z. Yordanova (✉)

University of National and World Economy, 8mi dekemvri, Sofia, Bulgaria

e-mail: zornitsayordanova@unwe.bg

have the ability to generate and spread conventional wisdom. Considerations about information systems' potential benefits in the field of new product development were suggested by Nambisan [3]. Raising the bar on the capability of information systems to facilitate business processes, product development, and innovation, we are analyzing their potential role for the organizational innovativeness (not organizational innovation only) in enterprises. Organizational innovation is proven to improve organizational performance and innovation performance [4]. It needs further to be digitalized because of the core principles of digital transformation based on a continuous improvement cycle, involving innovative and digital capabilities of an enterprise [5]. To overcome some of the main shortcomings of a team's innovation process in enterprise modeling and many innovation failures, here are formulated three research questions (RQ), the answers to which motivate this case study research.

RQ1: How does a single-team innovation approach can be transformed into an organizational innovation model for a whole enterprise?

RQ2: What are the pillars for increasing organizational innovation throughout the organization, regardless of the skills of innovative talent?

RQ3: What are the needs for digitization of a simple innovation process, easy to understand for all participants in innovation in an enterprise?

RQ4: How can an organization motivate employees to share their ideas, problems, goals, and expertise for organizational innovation? We answer these by analyzing a case study with the platform GAPideaz—an enterprise Web system for organizing and digitizing a simple innovation process encompassing all employees of an enterprise and focusing team efforts for the purpose of innovation development and problem solving.

2 Theoretical Background

2.1 *Organizational Innovation*

Organizational innovativeness is considered from different points of view in the literature and has been in the focus of scientists for a long time. It had been scoped as the relationship between organizational determinants of innovation, types of innovations, and measures of organizational performance by Subramanian and Nilakanta back in 1996 [4]. At the beginning of the theory development, it was seen only as a means for strengthening product development and re-organizing strategies for better commercialization of innovations [6]. Organizational innovativeness has later been seen as organizational effectiveness of innovation and a framework based on a model of organizational innovation grounded on key concepts found in literature on innovation was proposed by Tang [7]. These six KPIs are (1) project raising and doing, (2) knowledge and skills, (3) behavior and integration, (4) information and communication, (5) guidance and support, and (6) the external environment. Some other

studies analyzed organizational innovation as a complex system handling organizational culture, knowledge sharing, and organizational innovation [8]. Breaking down organizational culture for innovation, 12 organizational culture values supporting innovation have been identified backed by 33 cultural dimensions promoting an innovation culture [9]. Recent research has established a link between technology and innovation in the organization [10], which further raises the question of how to organize innovation occurring in enterprises. The enterprise ability to handle organizational issues associated with the innovation process are possible mechanisms through which organization capital affects firm innovation positively [11].

2.2 Digitalization of Innovation Development

Several calls for digitalization of innovation development and innovation processes in enterprises have been raised during the last 20 years. As innovative organizations rely on multiple sources for ideas and on multiple channels and mechanisms to bring ideas to further development [7], digitalization may answer to this gap in the enterprise information systems excluding innovation management as a holistic management approach. Some modern aspects and problem areas in the field of digitalization of the innovation process were raised by Bykovskaya et al. [12] who emphasized on the lack of prominent information technologies addressing the enterprise innovation processes encompassing all functional areas of innovation management across an organization. Digitalization of innovation management in the literature focuses mostly on platforms, toolkits, and communities that have been established and designed to generate ideas for innovation [13]. Some researchers on innovation management even concluded that digitalization of innovation management would lead to its reinvention and total transformation [14]. Digitalization of innovation management often overlaps with digital transformation. Thus, it could be theoretically scoped with encompassing all kinds of combined effects of several digital innovations bringing about novel actors (incl. stakeholders, systems, organizational set ups), structures, practices, values, and beliefs that change, threaten, replace, or complement the status quo within organizations. Hinings et al. [15] identified three types of novel institutional arrangements, critical for digital transformation: digital organizational forms, digital institutional infrastructures, and digital institutional building blocks which all are also much relevant for digitalization of innovation.

2.3 Enterprise Systems for Innovation Management

As Swanson [1] suggested, the innovation management as part of the operating enterprise systems is based on information flow and this information consistency and

smoothly spreading. This study distinguished three types of potential innovation-driven information systems: (1) innovation mechanisms embedded into the information system tasks; (2) innovation supported administration of the business; and (3) innovations embedded in the core technology. Further, information systems have been analyzed for automating and facilitating innovation processes and assessments of innovation outcomes by Nambisan [16]. The scholar proposed in his article a framework for better understanding of IT's role as a powerful resource for innovation development and management. Limited empirical proof motivated a call for further research by Kyriakou et al. [17] concerning the effects of the many different types of enterprise systems that firm use for their innovation performance. Their study investigated empirically and compared the effects of six important and widely used types of management information systems (ERP, CRM, e-sales, telework, and collaboration support systems) on firms' product, service, and process innovation. All these information systems showed some positive effects on both product/service and process innovation during the experimentation, especially the e-sales systems and CRM on product/service innovation. Another research on information systems and their possible impact over the innovativeness of an organization was conducted as an empirical research by focusing on eight criteria summarizing innovation pillars [18]. These were identified to be (1) organizational experience accumulation; (2) expanding motivation; (3) systematization and storage of learned lessons from successful projects; (4) talents' characteristics; (5) creativity management (ability to generate ideas, critical thinking, and creative problem solving); (6) change management organizational capabilities; (7) integration of perspectives into the business processes (openness to ideas, research orientation, cooperation); (8) entrepreneurship (identifying problems, seeking improvement, gathering information, independent thinking, technological consciousness, result orientation).

Besides the established practice that information systems, especially those supporting business processes, benefit the innovation development and management, some research showed that management information systems may affect the deployment and use of technological innovations [19] which could indirectly reinforce organizational innovation. Such tools enable the exchange of experiences and information transfers to the management levels to sustain competitive advantage and support decision-making [20], critical for innovation management, and accumulating organizational innovativeness. Even more, organizations that do not have information systems and processes for formal sharing practices would risk to fail in leveraging their managers' intellectual capital for business innovation and growth [21]. Going further, operational efficiency as an element of organizational innovativeness has also been seen in the literature as a goal of enterprise systems for innovation [22].

3 Research Design

We adopted a single-case study in order to in-deeply observe and analyze how digitalization of innovation management across an organization may engage innovation

talents and thus to support the establishment and digitalization of organizational innovativeness. As Zainal suggested, we select this research approach to explore and understand the complex issues related to organizational innovativeness and go beyond the quantitative results and understand the behavioral conditions through organizational perspective [23]. For addressing the research questions, the GAPideaz platform is observed and critically analyzed in the result and discussion's section with relevant analyzes about how it manages and digitizes the innovation process and creates a foundation for building organizational innovativeness. During the research, action research methods had taken place and the author met and discussed the obtained results from using the platform in 2 enterprises. The author of the research collaborated with the creators of the platform and some of the platform's users to bring more insights on its operation and impact on processes and users behavior for innovation as well as on encompassing these into a single platform. The presented results are downgraded to the level of marketing explanation for potential users since the case study aims at contributing mostly for non-innovation professionals, enterprise owners, employees from non-product development departments, and also to scholars in general innovation and technology management.

4 Results and Discussion

In this section of the study, we present the GAPideaz solution (<https://www.gapideaz.com/>) and discuss this example of a platform for establishing organizational innovation and managing it in an enterprise. In the same section, we focus on the research questions we asked at the beginning of the study and answer them.

4.1 GAPideaz Case Study

GAPideaz is a corporate platform for organizing innovation efforts for and by all employees and members of an organization. It has been seen as a corporate booster that digitalizes into a simple innovation process many activities such as idea generation, problem definition, call for action, goal setting, assessments, expertise sharing, project management tools for building ideas into projects for innovation, educational blog section. The concept behind the platform is that every employee, no matter of the rank, position, and experience has its own profile in the platform. The platform is Web-based and allows employees to express their innovativeness 24/7. Every employee can write down their ideas, problems, goals, and also to provide their own expertise for the benefit of the others' ideas, problems, and goals. All this information is public, but anonymous, unless a project is created. In projects, there are common resources, employees associated with the tasks, statuses of phases and all kinds of functionalities, necessary for innovation development. However, for kicking off an innovation project, all pre-project phases are facilitated through the platform

GAPideaz. Examining the platform, the main pillars of it are ideas, problems, goals, expertise and education for innovation culture and sharing. The platform maintains diverse functionalities facilitating collaboration such as association/submission of ideas for solving problems and achieving goals, automation of all applications of ideas for the owner of a problem or a goal, project management organization with sprint and tasks distribution as well as task status updates.

4.2 Organizational Innovativeness through GAPideaz

A main value of the innovation process designed by GAPideaz platform is the innovation process’ openness across the enterprise. It is a common practice for an innovation, an improvement or optimization to take place in a single team/department instead of reflecting to the whole organization. Even more inefficiently is the practice in many organizations to perform these innovation-related activities by single teams instead of engaging all employees from the organization in the process. Involving all employees especially in the first phases of the innovation process may only bring more ideas, clearer problem definition, and scalable solutions relevant for more users. However, such employees’ engagement in the organization’s problem solving and innovation is hard for managing. This is why, in this study, we use the case study approach in order to elicit the design elements and assumptions behind GAPideaz and discuss some results that the platform brings to users.

The results of a single-team innovation development resemble a marketing funnel or a closed innovation development model where all kinds of triggers for innovation such as problems, new ideas, goals, projects, team work, and regulations are put into a funnel which funnel aims at processing and transforming them into an innovation.

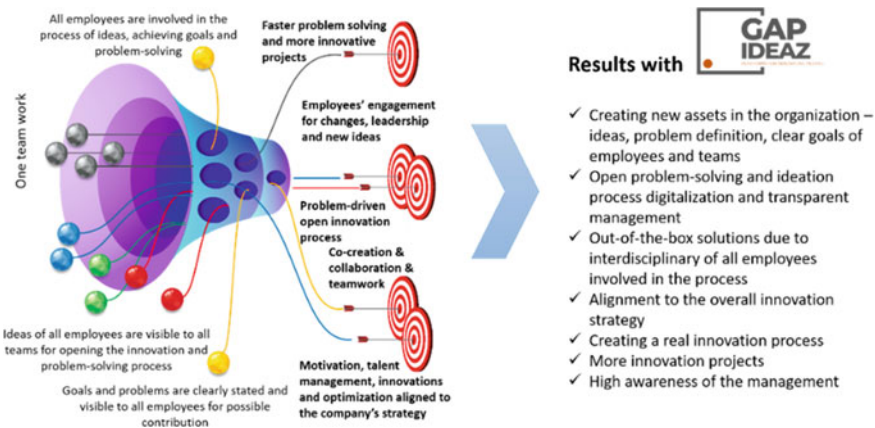


Fig. 1 Multi-actors’ innovation development with GAPideaz

On Fig. 1, we present a funnel for innovation development, but in this very case, it is open for additional collaboration from more employees and teams by the organization. This can only bring more innovation-related ingredients. The discussion on whether involving more actors in the innovation process could bring more positive results or would rather complicate it is not in the scope of this case study. Here, we are examining a process and a digitalization tool for those organizations which have already reached the usual difficulties and shortcomings of a closed innovation process, locked into single teams. The benefits of opening the internal innovation process by making it public for the members of an organization could boost the alignment of employees' activities, ideas, and problem solving to the overall innovation strategy could also increase the awareness of the management for the innovation attitudes among the employees and vice versa. The summarized outcome of using such a platform is organizing the efforts of all employees for the benefit and goals of the organization. The platform may also facilitate some human resource management tasks such as detecting talents, motivating and engaging employees, facilitating collaboration, and communications for complex tasks. The need for digitalization of innovation-related tasks of employees comes from a few factors that affect the organizational innovativeness. These factors have increased their significance for the organizations in 2020 and 2021 when a big share of them apply home office practice (as a consequence of COVID-19). Major reason for digitalization is the fact that creativity and innovativeness of people do not show only in the working hours and they need a tool to express these 24/7. Another cause of the need for digitalization of innovation processes across organizations is the impossibility for effective and intensive collaboration between more than 30–50 people. This factor is extremely relevant for medium and large organizations with more than 100 employees, working in different departments, locations, roles who do not communicate frequently in regard to their daily operations. The need for sharing ideas, problems, engagement, expertise, and common goals have been largely discussed in the context of organizational innovation and leading creative people [24]. This is how, together with GAPideaz solution's founders and feedback from their clients, we summarized a possible solution for these shortcomings. The next figure presents the ordinary situation in enterprises when it comes to employees' sharing (according to some feedback of GAPideaz's customers) and the resolution after implementing the platform. Innovation mostly relies on team working and collaboration [25] and digitalizing innovation processes involve not only automation of activities, but substantially digitalizing the communication between stakeholders and potential actors of the process. During the case study examination, GAPideaz creators shared primary data of the previous research stating that most of 80% of the employees are too shy to proactively participate in innovation processes, to share their opinion, problems needed solutions, ideas, or even expertise.

Figure 2 shared the results of digitalization innovation process through GAPideaz, overcoming the major barrier for it—communication and proactiveness bottlenecks.

How do employees share ideas for change, optimization, problem-solving and innovation with GAPideaz?

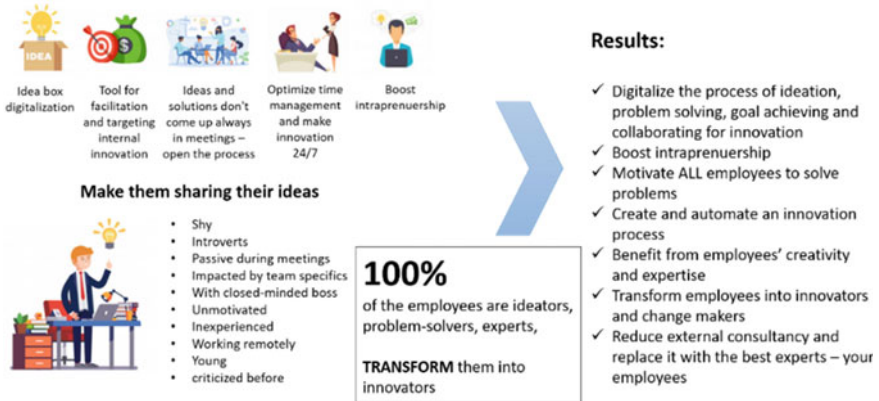


Fig. 2 Overcoming the process of sharing ideas, problems, solutions, expertize with GAPideaz platform (based on feedback of users)

4.3 Insights for Organizing and Digitalizing Innovation across Enterprise

In this section, the study addresses the research questions for outlining the general outcomes of digitalization of innovation process and organizational innovativeness in organizations through the design of GAPideaz platform.

RQ1: How does a single-team innovation approach can be transformed into an organizational innovation model for the whole enterprise? The discussion on this question was developed in the previous section of the paper and reveals some common assumptions of the single-team innovation efforts which the design of the tool and the digitalization element tries to overcome. Generally, the concept of open innovation is applied into an organizational design for the purposes of innovation development engaging and involving all employees.

RQ2: What are the pillars for increasing organizational innovation throughout the organization, regardless of the skills of innovative talent? According to the design of the analyzed tool, the pillars for organizing innovation efforts across the organization (bearing in mind the involvement of non-professional innovators), are ideas (idea generation, idea search, idea observation, idea register, idea assessment), goals and problems (problem definition, goal achievement), expertise sharing, education on innovation through gamified exercise and tools, project management functionalities for bringing ideas, problems, and goals into innovation projects.

RQ3: What are the needs for digitization of a simple innovation process, easy to understand for all participants in innovation in an enterprise? Altogether, the digitalization value comes from (1) remote working, personal behavior specifics

for (2) being creative 24/7 and the ability for organizations to take advantage from their employees creativity and innovativeness, (3) uncertainty and self-doubt of employees to bring public their ideas and problems, (4) the spaghetti designed communication and inter-departments channels for innovation across organizations, (5) the lack of engagement tool for literally anybody in the organization to be innovator.

RQ4: How can an organization motivate employees to share their ideas, problems, goals, and expertise for organizational innovation? According to GAPideaz customers, transparency and digitalization of the process into clear and simple steps are the strongest motivational factors for employees to participate in enterprise innovation processes.

5 Conclusion

In conclusion, we can summarize that digitalization of innovation efforts and engaging all employees into the firm's innovation processes of an organization may bring many new benefits which are not currently addressed by the traditional means of organizational design or innovation strategizing and enterprise modeling for innovation. The examined design for organizational innovativeness in this paper borrows some elements from open and user innovation, process management, organizational innovation, and digitalization for revealing a possible process for innovation development with the following pillars: ideas, goals, problems, expertise, and education for culture of innovation. The implementation and application of the researched tool GAPideaz may facilitate diverse organizational issues related to employees' motivation and engagement, problem-driven innovations, registering the ideas of employees, and bridging the gap between management and employees when it comes to innovation management. Major insight to the innovation management theory is the resulted outcome of the case study that a factor that employees value the most in their engagement is the transparency of the innovation process for all employees with easy and clear public access to such a digital tool.

Acknowledgments The work on the study and publication was supported by the UNWE NID 9/2021.

References

1. Swanson, E.B.: Information systems innovation among organizations. *Manage. Sci.* **40**(9), 1069–1092 (1994)
2. Baskerville, R.L., Myers, M.D.: Information systems as a reference discipline. *MIS Q.* **26**, 1 (2002)
3. Nambisan, S.: Information systems as a reference discipline for new product development. *MIS Q.* **27**(1), 1–18 (2003). <https://doi.org/10.2307/30036517>

4. Subramanian, A., Nilakanta, S.: Organizational innovativeness: exploring the relationship between organizational determinants of innovation, types of innovations, and measures of organizational performance. *Omega* **24**(6), 631–647 (1996)
5. von Leipzig, T., Gamp, M., Manz, D., Schöttle, K., Ohlhausen, P., Oosthuizen, G., Palm, D., von Leipzig, K.: Initialising customer-orientated digital transformation in enterprises. *Procedia Manuf.* **8**, 517–524 (2017)
6. Normann, P.: Organizational innovativeness: product variation and reorientation. *Adm. Sci. Q.* **16**(2), 203–215 (1971)
7. Tang, H.K.: An inventory of organizational innovativeness. *Technovation* **19**(1), 41–51 (1998). ISSN 0166-4972. [https://doi.org/10.1016/S0166-4972\(98\)00077-7](https://doi.org/10.1016/S0166-4972(98)00077-7)
8. Azeem, M., Ahmed, M., Haider, S., Sajjad, M.: Expanding competitive advantage through organizational culture, knowledge sharing and organizational innovation. *Technol. Soc.* **66**, 101635 (2021)
9. Kashan, A.J., Wiewiora, A., Mohannak, K.: Unpacking organisational culture for innovation in Australian mining industry. *Resources Policy* **73**, 102149 (2021)
10. Donbesuur, F., Ampong, G.O.A., Owusu-Yirenkyi, D., Chu, I.: Technological innovation, organizational innovation and international performance of SMEs: the moderating role of domestic institutional environment. *Technol. Forecast. Soc. Change* **161**, 120252 (2020)
11. Li, Y., Song, Y., Wang, J., Li, C.: Intellectual capital knowledge sharing, and innovation performance: evidence from the Chinese construction industry. *Sustainability* **11**(9), 1–20 (2019)
12. Bykovskaya, E.N., Kharchilava, G.P., Kafiyatullina, Yu.N.: Modern trends of digitalization of innovation process. *Upravlenie* **6**(1), 38–43 (2018). <https://doi.org/10.26425/23093633-2018-1-38-43>
13. Dong, J.Q., Wu, W.: Business value of social media technologies: evidence from online user innovation communities. *J. Strateg. Inf. Syst.* **24**(2), 113–127 (2015)
14. Nambisan, S., Lyytinen, K., Majchrzak, A., Song, M.: Digital innovation management: reinventing innovation management research in a digital world. *Manage. Inf. Syst. Q.* **41**, 223–238 (2017)
15. Hinings, B., Gegenhuber, T., Greenwood, R.: Digital innovation and transformation: an institutional perspective. *Inf. Organ.* **28**, 52–61 (2018)
16. Nambisan, S.: Information technology and product/service innovation: a brief assessment and some suggestions for future research. *J. Assoc. Inf. Syst.* **14**(4), (2013). Article 1. <https://doi.org/10.17705/1jais.00327>
17. Kyriakou, N., Loukis, E., Arvanitis, S.: Enterprise systems and innovation—an empirical investigation. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 4686–4696 (2016). <https://doi.org/10.1109/HICSS.2016.581>
18. Yordanova, Z.: Management information systems (MIS) boost organizational innovation competencies. In: International Conference on Application of Information and Communication Technology and Statistics and Economy and Education (ICAICTSEE) (2018)
19. Jamwal, D., Singh, N.: Management information system competencies and strategies. *IJCST* **2**(1), pp. 109–112, (2011). ISSN: 2229-4333(Print) | ISSN: 0976-8491(Online)
20. Karim, A.G.: The significance of management information systems for enhancing strategic and tactical planning. *JISTEM J. Inf. Syst. Technol. Manage.* **8**(2), (2011)
21. O’Neill, B.S., Adya, M.: Knowledge sharing and the psychological contract: managing knowledge workers across different stages of employment. *J. Manage. Psychol.* **22**(1), 411–436 (2007). <https://doi.org/10.1108/02683940710745969>
22. Bhattacharya, P.J., Seddon, P.B., Scheepers, R.: Enterprise systems for innovation in products and processes: beyond operational efficiency. In: 23rd Australasian Conference on Information Systems Enterprise Systems for Innovation 3–5 Dec 2012, Geelong (2012)
23. Zainal, Z.: Case study as a research method. *Jurnal Kemanusiaan bil.* **9**, (2007)

24. Mumford, M.D., Scott, G.M., Gaddis, B., Strange, J.M.: Leading creative people: orchestrating expertise and relationships. *Leadersh. Q.* **13**(6), 705–750 (2002)
25. Zahoor, N., Al-Tabbaa, O.: Inter-organizational collaboration and SMEs' innovation: a systematic review and future research directions. *Scand. J. Manage.* **36**(2), 101109 (2020). <https://doi.org/10.1016/j.scaman.2020.101109>

What Changed in One Year of a Pandemic and What the Portuguese are not Willing to Admit: Consumer Neuroscience and Predictive Analytic Contributes to Communication Strategy



Valentina Chkoniya , Dorota Reykowska, Rafal Ohme, and Ana Côrte-Real

Abstract The world is headed towards a new normal, while pandemic is continuing with new waves hitting nations. However, this storm will pass, but the choices we make now could change our lives for years to come. It is crucial to use new research perspectives bringing scientists and practitioners together to support decisions that would shape our future. This paper gives an example of consumer neuroscience and predictive analytics helping to go beyond rational verbatims to investigate real deep nonconscious convictions that people may not even be fully aware of and which cannot be covered by traditional opinion surveys. For example, public confidence in the healthcare system is related to pro-social behaviour and compliance with non-pharmaceutical interventions during a crisis. The success in controlling pandemics depends on behaviour, and health officials need to persuade the population to make behaviour changes to ensure success. By providing comparative results from Portugal, this study highlights what changed in one year and what the population is not willing to admit by using the “COVID-19 Fever” project data collected and analysed with the iCode Smart Test in 2020 and 2021 and offers valuable data to support an effective communication strategy.

Keywords COVID-19 · Pandemic · Consumer neuroscience · Advanced analytics · Consumer behaviour · iCode · Response time testing

V. Chkoniya (✉)
GOVCOPP, ISCA-UA, University of Aveiro, Aveiro, Portugal
e-mail: valentina.chkoniya@ua.pt

D. Reykowska · R. Ohme
NEUROHM, Warsaw, Poland
e-mail: dorota.reykowska@neurohm.com

R. Ohme
e-mail: ralfal@neurohm.com

R. Ohme
WSB, Warsaw, Poland

A. Côrte-Real
Católica Porto Business School, Porto, Portugal
e-mail: acortereal@ucp.pt

1 Introduction

Every crisis is an opportunity to reflect on deeper social issues and improve matters for the long run. It is crucial to use new research perspectives bringing scientists and practitioners together to support decisions that would shape our future, knowing that this storm will pass, but the choices we make now could change our lives for years to come [1, 2]. Behavioural research needs to be stepped up so that challenges facing humankind can be taken care of better. Consumer neuroscience (CN) is a branch of research still under development, and it required a form of technology transformation as traditional researchers and marketers came to accept that nonconscious response and emotion are critical in driving perceptions and behaviours. This budding revolution was also enabled by scientific advancements in predictive analytics (PA) [3]. This paper gives an example of CN and PA that help to go beyond rational verbalisms to investigate real deep nonconscious convictions that people may not even be fully aware of and which cannot be covered or even misleads by traditional opinion surveys. By providing comparative results from Portugal, this study highlights what changed in one year and what the population is not willing to admit by using the “COVID-19 Fever” project data collected and analysed with the iCode Smart Test in 2020 and 2021. The data reflects how everyday behaviours in Portugal have been transformed in a way that may be here to stay. Some of the changes are highly visible, such as a relaxed attitude about complying with the recommendations for physical distancing, but an analysis of different attitudes during one year of pandemic flags up other adjustments we may not be so aware of, such as the fact that Portuguese are more open to helping more vulnerable people and grooving concern about the healthcare system. Insights outlined in this paper can help to support an effective communication strategy and may also be relevant to future public health crises.

2 Background and Related Work

2.1 *Behaviour is the Key to Success*

The success in controlling pandemics depends on behaviour, and health officials often need to persuade the population to make several behaviours changes [4–6]. A century ago, Science magazine published the research on the Spanish Flu pandemic. The paper underlines three main lessons learned from it [4, 7]:

1. people do not appreciate the risks they run
2. it goes against human nature for people to shut themselves up in rigid isolation as a means of protecting others
3. people often unconsciously act as a continuing danger to themselves and others.

Declarations and actions do not always overlap, and thus, predicting future behaviour solely based on self-reported measures seems to be ineffective [8]. **Where**

are these discrepancies coming from? Social expectations, conformity, trying to make a good impression, or trying to look smart. All these can impact how people answer [1]. The behaviour of individuals living in communities is regulated by moral norms and values [9–11]. People who do what is “right” are respected and publicly admired, while those who do what is “wrong” are devalued and socially excluded [4, 12]. The relationship between trust, confidence, and compliance is a crucial element in human behaviour. Fostering this connection is a reciprocal task of authorities and citizens—one requiring future engagement foundation of trust and reciprocal cooperation, which is an investment that pays dividends during future crises [13, 14]. The World Health Organization reports highlighting the connection between communication and behaviour stating that “communication is seen to have relevance for virtually every aspect of health and well-being, including disease prevention, health promotion and quality of life” [15].

2.2 *Communication Strategy*

Communication strategy is the scheme of planning how to share information effectively [16]. The information environment around a pandemic underscores the importance of effective communication [4].

Enlisting trusted voices has been shown to make public health messages more effective in changing behaviour [17]. It is common for trust, in general, to deteriorate in situations of high uncertainty and ambiguous behaviour [18]. The information environment around a pandemic underscores the importance of effective science communication [4]. The results of the study highlight pathways for the adjustments in the decision support systems and strategic planning for it. Decades of research have found that, whether recipients are motivated to think carefully or not, sources perceived as credible are more persuasive. The credibility of sources stems from how trustworthy and expert they are perceived to be [4, 19–21].

After identifying the credible source, it is important to identify, what message should be delivered. Different messaging approaches can be efficient, including emphasizing the benefits to the recipient, focusing on protecting others (e.g. “wash your hands to protect your grandparents”), aligning with the recipient’s moral values, appealing to social consensus or scientific norms, and highlighting social group approval [4, 16, 22–27]. To slow down the spread of the pandemic, most countries attempted to implement various social distancing measures and other non-pharmaceutical interventions (e.g. bans on large gatherings, closure of schools, gyms, bars, and restaurants, and mandating stay-at-home policies). The communication strategy of these measures was mainly based on a combination of extrinsic motivations such as social pressure, intrinsic motivations that include moral support and social norms, and is dependent on factors such as local income, risk-taking behaviour, personality characteristics, political orientation, trust in the government, trust in media sources, and belief in science [13, 28, 29]. **Which of these messages work best?** CN and PA become the support for this type of decision [8, 16, 29]. With

so many changes within a short period comes an inevitable psychological impact. The communication message during lockdown was accepted because it was an exceptional situation, and the subsequent adjustment to new habits has been less well tolerated because when people return to the former situation but cannot do what we did before, they find it more difficult to adapt. The response time (RT) measurement offers results that are less influenced by self-presentation bias or social desirability than traditional survey [30–32].

3 Methodology

3.1 Research Method

To uncover real attitudes about the pandemic, iCode smart test was used. iCode is a Web-based and device-agnostic technology created by NEUROHM with RT testing (RTT) tool based on Fazio’s attitude accessibility paradigm. According to this model, attitudes can be explained by two factors (1) explicit, declared opinion and (2) implicit, accessibility of attitudes. Opinions that are strong and accessible are expressed with high confidence and indicated by faster RT, while weaker, less accessible attitudes expressed with hesitation are indicated by slower RTs [33]. Fazio has shown that correlations between attitudes and behaviour are higher among people with fast RT when expressing their opinions, which means strong attitudes have a more significant influence on behaviour [34].

This research builds on “COVID-19 Fever” international project that has emerged to better understand people’s attitudes and opinions regarding the pandemic. It aims to assist and inspire effective communication by the health and media authorities, as well as to monitor the progress of the “COVID-19 Recovery” process. Because the crisis requires large-scale behaviour change and places significant psychological burdens on individuals, insights from the social and behavioural sciences can be used to help align human behaviour with the recommendations [4]. This paper focuses on comparing the results of 2 waves of research done in Portugal. Wave 1 was conducted during the first lockdown period in Portugal April–May 2020. Wave 2 was conducted a year later. A detailed description of the results from Wave 1 was presented in the previous ICMarTech conference in Lisbon [1].

3.2 Process

The study was performed online in two waves: Wave 1 in April–May 2020 gathering $n = 291$ Portuguese citizens and Wave 2 in May 2021 with $n = 243$ Portuguese citizens. All the tasks were presented on a computer screen. The main part of the test was focused on gathering people’s opinions on the COVID-19 pandemic. We have

tested ten statements about life during the pandemic, grouped into five categories (two statements each): *worries about self*, *worries about others*, *trust in the healthcare system*, *following guidelines*, and *pro-social behaviour* (Table 1).

In comparison with *Wave 1*, a change was made to the *following guidelines* category. The statement: “*I comply with the restrictions to stay home*” was substituted by “*I wash hands for 20 s when necessary*”. The reason for this change was connected with the fact that during *Wave 2* of the research there was no lockdown in Portugal, so the initial statement was irrelevant.

The task of the participants was to evaluate if they agree with the presented statements. The statements were shown individually, and the answers were given on a three-point scale: *yes*, *hard tell*, *no*. Participant’s declarative responses, as well as the time needed to answer (RT testing), were recorded.

To make sure that the registered RT for both *yes* and *no* answers is not biased, that is influenced by the position of the mouse, a control screen (Advanced Button) was introduced. After the answer to each statement was given, a whiteboard, with a blue button in the middle, was presented. The task of the respondents was to click on the blue button. This ensured that each time, the distance to both *yes* and *no* buttons was the same.

The test questions have been preceded by a calibration phase. This phase was focused on achieving three goals:

- a. Familiarization with the scale
- b. Familiarization with the purpose of the study and the type of the tasks
- c. Increasing participant’s focus on the task.

Table 1 Statements used in the study

Category	Statement
Following guidelines	I comply with the recommendations for physical distancing
	I wash hands for 20 s when necessary
Pro-social behaviour	I actively encourage others to follow the restrictions and guidelines
	I would like to help people who are more vulnerable to COVID-19
Evaluation of health care system	I am satisfied with how our healthcare system is handling this crisis
	In case of a coronavirus infection, I will get appropriate medical help
Worries about others	I am worried about the health of my older family members
	I am worried about the health of people in my country
Worries about self	I’m worried about my financial situation
	I am worried about my health

3.3 Analysis

The first step of the analysis was to clean the data. Values lower than 500 ms and higher than 10,000 ms were eliminated. Additionally, respondents who were very quick or very slow in providing their answers in over 40% of responses have been excluded from the analysis as people who did not pay enough attention to the task. During the analysis, the milliseconds were converted to natural logarithms [35], and the extreme values were truncated. The next step was to standardize the individual values into a scale from 0 to 1000, with the bottom values indicating the slowest RTs and the top values representing the shortest RTs.

The last step was to select the fast responses, indicating a confident answer. A cut-off value of 600 has been taken as a threshold above which the results were treated as fast response. The proposed value has been calculated based on 15 studies, including approximately 3500 respondents, and represents a unified value (benchmark) above which responses were significantly higher than average. In the following sections, this analysis will be referred to as *confident yes/no* answers.

The analysis is focused on two measures: *declarative* answers (explicit) and *confident* answers (a measure of explicit response enriched with the implicit score, based on RT). Explicit and implicit scores are not treated as contradictory, and the implicit measure is used to supplement and deepen the explicit results.

Additionally, the differences between declarative answers and confident answers (*gap ratios*) were introduced. This is the % of people who expressed hesitation when answering (out of all who said *yes*) and was calculated using the formula below.

$$\text{Gap ratio} = ((\text{yes} - \text{confident yes})/\text{yes}) * 100\%$$

4 Results

Continuing the approach taken in the analysis of *Wave 1*, we will be focusing only on *yes* answers, as in the study, there is a significant dominance of affirmative answers. For *Wave 1*, in 4 out of 5 tested categories, at least 70% of people agreed with the tested statements. For *Wave 2*, it was just 3 out of 5 categories with at least 70% agreement.

People declare that they are concerned about the *health of others* (88%), followed closely by pro-social behaviour (79%), 75% declare *following guidelines*, and 66% positively evaluate the *healthcare system*. The lowest number of affirmative answers was observed for *worries about self* (57%) (see Fig. 1).

Analysis of the merged explicit and implicit answers (*confident yes*) shifts the ranking of categories. The confidence for *worries about others* and *following guidelines* is the highest (62% and 54%, respectively) followed by *worries about self* and

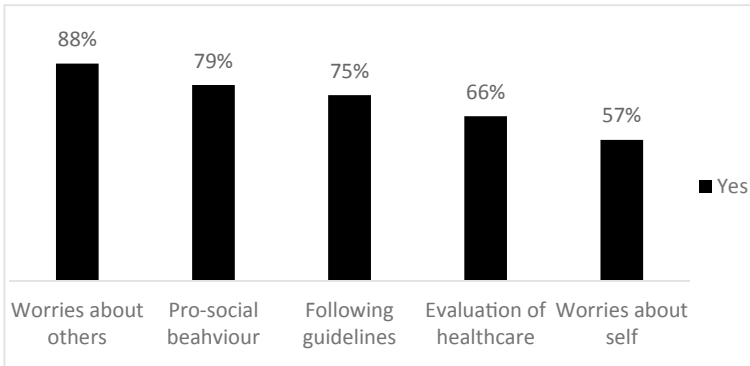


Fig. 1 Explicit answers to statements in each category (Wave 2) (Per cent of people who explicitly agree with the statements)

pro-social behaviour (both at 42%). The confidence scores of the satisfaction with the *healthcare system* are the weakest 13% (see Fig. 2).

The significant discrepancy between *yes* and *confident yes* answers is reflected by high scores in the gap ratio. The highest numbers were obtained for two categories: readiness to perform *pro-social behaviour* (47%) and *evaluations of the healthcare system* (80%). For the other three categories, the gap ratio oscillates between 27 and 29% (see Fig. 3).

The obtained trends are in line with the results from *Wave 1*, where we have also observed the highest tendency to distort explicit declarations in two categories *pro-social behaviour* (62%) and *evaluations of the healthcare system* (61%).

WAVE 1 versus WAVE 2 Have the attitudes of people changed? The explicit and implicit results from Wave 1 and Wave 2 were compared using the z-test for independent-sample proportions. Significant results at $p < 0.05$ were marked on the

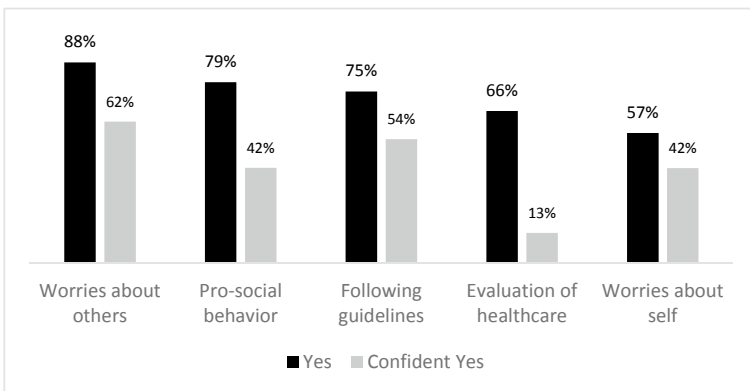


Fig. 2 Explicit and implicit answers to statements in each category (Wave 2)

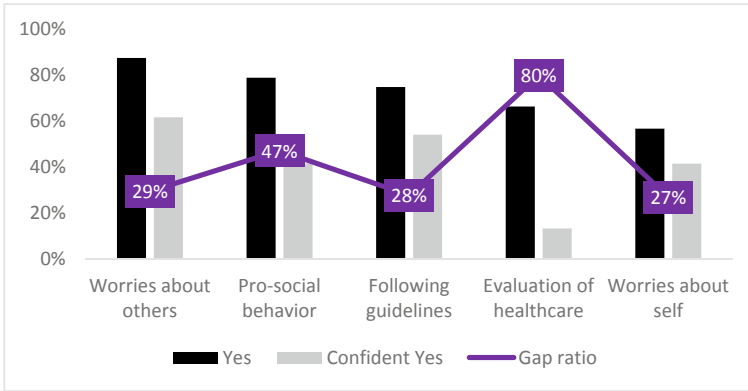


Fig. 3 Gap ratio for the explicit vs merged explicit and implicit answers

graphs with arrows. Dotted arrows represent difference at a tendency level $p < 0.1$. When comparing the results from the two waves on the declarative level, we observe a significant decrease in Wave 2 for two categories *worries about others* and *following guidelines* (see Fig. 4).

When analysing the changes in confident answers, as observed in the explicit results, in *Wave 2*, people are less confident about *following guidelines* and less *worried about others* (difference at a tendency level $p < 0.1$). The analysis of the confident answers brings, however, an additional layer. What shows up is a significant *drop in the evaluation of health care* and a significant *increase in pro-social behaviour* (Fig. 5).

The comparison of gap ratio between waves explains well the discrepancies between explicit and implicit results (Fig. 6). To compare the gap results, Mann–Whitney U test was used. People distorted their answers more in Wave 2 when it

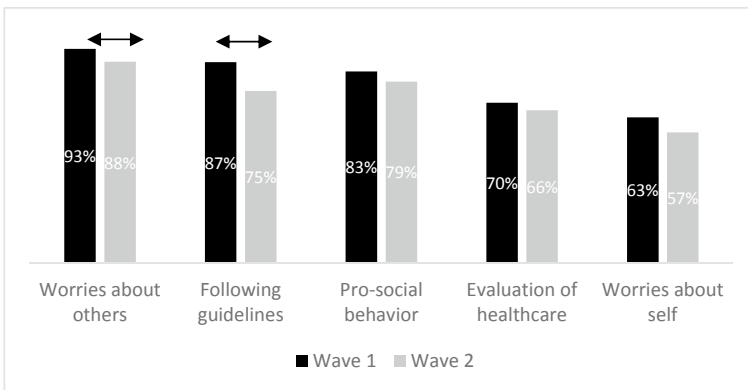


Fig. 4 Explicit answers to statements in each category (Wave 1 versus Wave 2)

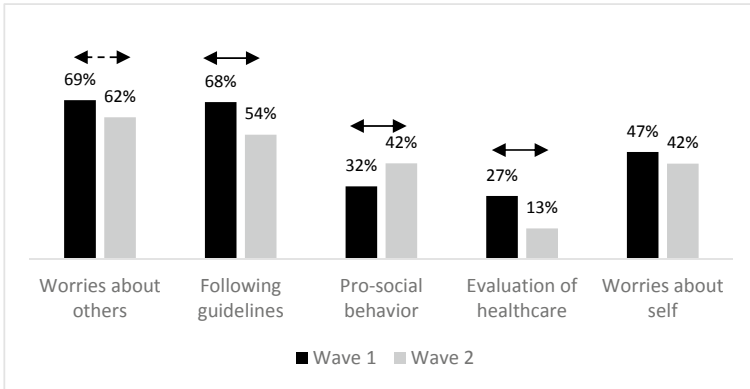


Fig. 5 Implicit answers to statements in each category (Wave 1 versus Wave 2)

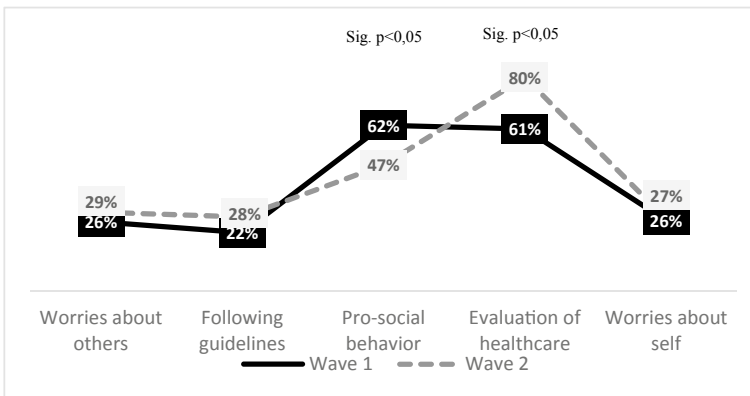


Fig. 6 Gap ratio per category. Comparison of waves

comes to *evaluating health care*. So, on the declarative level, we did not notice any changes between waves, while the confidence showed a significant drop—peoples’ positive opinions decreased, but they did not want to admit it.

On the other hand, the tendency to distort decreased for *pro-social behaviour*. So again on the declarative level, we did not notice any differences between waves, but the confidence revealed that after a year of the pandemic people are more *willing to help others* than at the beginning of the pandemic.

Portuguese continue to worry about others and trying not to show concerns about themselves. Feeling more relaxed about complying with the recommendations for physical distancing. People become more convinced about active pro-social behaviour and less confident about the healthcare system. It can be expected that the Portuguese are more open to helping more vulnerable people.

Confidence in the healthcare system implies an expectation that sufficient and appropriate treatments will be provided if needed. The COVID-19 public health crisis is a significant test of the behavioural implications arising from this confidence [13, 28]. In 2020, on a declarative level, Portuguese very positively evaluated the healthcare system, but only 27% were confident about their answers. One year later, in 2021, the declarative-level response remained almost the same, when confident responses drop down to 13%. This concern motivated people to have more active pro-social behaviour to encourage others to follow the restrictions as well as help those who are more vulnerable to COVID-19.

5 Conclusion

To mitigate the potentially devastating effects of the pandemic, action can be supported by the behavioural and social sciences [4]. From an exploratory perspective, thanks to PA, the results indicated that the RT score merged with the declarations score into a measure of confidence has significantly improved the ability to understand what is behind the motivation. Therefore, it can be a valuable expansion for traditional surveying techniques especially that the effort which is required to apply it is relatively low in comparison with other non-declarative techniques [36, 37]. The current study has demonstrated—through the use of a CN and PA—that public confidence in the healthcare system is related to pro-social behaviour during a pandemic crisis. Moreover, data reflects how the everyday behaviours of Portuguese have been transformed. Some of the changes are not detectable with a traditional survey (declarative level) and can be reviled only with the analysis of the confident answers. Another interesting aspect of the comparative analysis is the finding that low levels of healthcare confidence initially exhibit a more significant response and stay home, but during one year, this reaction plateaus more quickly and becomes to decline with the same magnitude. Less confidence in the healthcare system may signal citizens' awareness that it cannot handle an outbreak, leading to rapid and extensive self-isolation in the beginning. However over time with time lower confidence societies may “give up” on the stay-at-home measures but continue to comply with distancing requirements [13] and decreases over the year. Was observed a faster drop in compliance and increase in pro-social behaviour with the growth of the urgent need for health public services to be enhanced. Monitoring people's attitudes helps to trace change that requires constant narrative adjustment, empowered with new threads. Giving facts or scaring people all the time is not enough. When attitudinal change indicates that regarding education about the pandemic a glass ceiling was reached, new ways of spreading knowledge should be explored. For example, through humour, unsung heroes, or celebrities. Also, it is important to understand a moment when fear must give way to hope in the messages that are transmitted. Understanding the changes in people's attitudes and feelings during pandemics can help to identify the most credible source for each moment in time and define the most assertive message. Also, success in controlling pandemics depends on behaviour, and health officials need to persuade

the population to make behaviour changes to ensure success. The information environment around a pandemic is crucial for it underscores the importance of effective communication.

References

1. Chkoniya, V., Reykowska, D., Ohme, R.: What do Portuguese really feel about the COVID-19 pandemic CN and PA—discovering new perspectives BT. In: Rocha, A., et al. (2021) *Marketing and Smart Technologies*, pp. 283–294. Springer Singapore (2021).
2. Harari, Y.N.: Disease in a world without a Leade. *TIME Mag.* **195**, 46–47 (2020)
3. Moses, E., Clark, K.R.: *The Neuromarketing Revolution: Bringing Science and Technology to Marketing Insight*. IGI Global, (2020). <https://doi.org/10.4018/978-1-7998-3115-0.ch024>
4. Van Bavel, J.J., et al.: Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **4**, 460–471 (2020)
5. Dirks, N.: We need social science, not just medical science, to beat the pandemic. *Sci. Am.* (2021)
6. Margraf, J., Brailovskaia, J., Schneider, S.: Behavioral measures to fight COVID-19: an 8-country study of perceived usefulness, adherence and their predictors. *PLoS ONE* **15**, 1–22 (2020)
7. Soper, G.A.: The lessons of the pandemic. *Science* **49**, 501–506 (1919)
8. Ohme, R., Matukin, M., Wicher, P.: Merging Explicit Declarations With Implicit Response Time to Better Predict Behavior. IGI Global, (2020). <https://doi.org/10.4018/978-1-7998-3115-0.ch023>
9. Greene, J.D.: *Moral Tribes: Emotion, Reason, and the Gap*. Atlantic Books (2015)
10. Tomasello, M.: *Why We Cooperate*. xviii, pp. 206–xviii, 206 (2009)
11. Leach, C.W., Bilali, R., Pagliaro, S.: Groups and morality. In: *APA Handbook of Personality and Social Psychology, Volume 2: Group Processes*, pp. 123–149 (2015). <https://doi.org/10.1037/14342-005>
12. Haidt, J.: The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* **108**, 814 (2001)
13. Fai, H. et al.: How confidence in health care systems affects mobility and compliance during the COVID-19 pandemic. *PLoS ONE* (2020). <https://doi.org/10.1371/journal.pone.0240644>
14. O’Keefe, D.J., Hoeken, H.: Message design choices don’t make much difference to persuasiveness and can’t be counted on—not even when moderating conditions are specified. *Front. Psychol.* **12**, (2021)
15. Rimal, R.N., Lapinski, M.K.: Why health communication is important in public health. *Bull. World Health Organ.* **87**, 247–247a (2009)
16. Moses, E., Clark, K.R., Jacknis, N.J.: *The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience*, pp. 151–166. IGI Global (2021). <https://doi.org/10.4018/978-1-7998-6985-6.ch007>
17. Greyling, C., et al.: Lessons from the faith-driven response to the West Africa Ebola epidemic. *Rev. Faith Int. Aff.* **14**, 118–123 (2016)
18. Clots-Figuera, I., González, R.H., Kujal, P.: Trust and trustworthiness under information asymmetry and ambiguity. *Econ. Lett.* **147**, 168–170 (2016)
19. Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. In: *Communication and Persuasion*, pp. 1–24. Springer (1986)
20. Brinol, P., Petty, R.E.: Source factors in persuasion: a self-validation approach. *Eur. Rev. Soc. Psychol.* **20**, 49–96 (2009)
21. O’Keefe, D.J.: Message pretesting using perceived persuasiveness measures: reconsidering the correlational evidence. *Commun. Methods Meas.* **14**, 25–37 (2020)

22. Grant, A.M., Hofmann, D.A.: It's not all about me: motivating hand hygiene among health care professionals by focusing on patients. *Psychol. Sci.* **22**, 1494–1499 (2011)
23. Feinberg, M., Willer, R.: Moral reframing: a technique for effective and persuasive communication across political divides. *Soc. Personal. Psychol. Compass* **13**, e12501 (2019)
24. Lewandowsky, S., Gignac, G.E., Vaughan, S.: The pivotal role of perceived scientific consensus in acceptance of science. *Nat. Clim. Chang.* **3**, 399–404 (2013)
25. van der Linden, S., Leiserowitz, A., Maibach, E.: Scientific agreement can neutralize politicization of facts. *Nat. Hum. Behav.* **2**, 2–3 (2018)
26. Drummond, C., Fischhoff, B.: Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proc. Natl. Acad. Sci.* **114**, (2017)
27. Fainshtein, E., Serova, E.: Using Intelligent Text Analysis of Online Reviews to Determine the Main Factors of Restaurant Value Propositions. IGI Global (2021). <https://doi.org/10.4018/978-1-7998-6985-6.ch010>
28. Friedson, A.I., McNichols, D., Sabia, J.J., Dave, D.: Shelter-in-place orders and public health: evidence from California during the COVID-19 pandemic. *J. Policy Anal. Manage.* **40**, 258–283 (2021)
29. Bughin, J.R.J., Cincera, M., Reykowska, D., Ohme, R.: Big Data Is Decision Science: The Case of COVID-19 Vaccination. IGI Global (2021). <https://doi.org/10.4018/978-1-7998-6985-6.ch006>
30. Greenwald, A.G., Krieger, L.H.: Implicit bias: scientific foundations. *Calif. Law Rev.* **94**, 945–967 (2006)
31. Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., Banaji, M.R.: Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* **97**, 17–41 (2009)
32. Plangger, K. *Thriving in a New World Economy: Proceedings of the 2012 World Marketing Congress* (2016). <https://doi.org/10.1007/978-3-319-24148-7>
33. Fazio, R.H., Powell, M.C., Williams, C.J.: The role of attitude accessibility in the attitude-to-behavior process. *J. Consum. Res.* **16**, 280–288 (1989)
34. Fazio, R., Sanbonmatsu, D., Powell, M., Kardes, F.: On the automatic activation of attitudes. *J. Pers. Soc. Psychol.* **50**, 229–238 (1986)
35. Winer, B.J.: *Statistical Principles in Experimental Design*. McGraw-Hill (1971)
36. Ohme, R., Matukin, M., Pacula-Lesniak, B.: Biometric measures for interactive advertising research. *J. Interact. Advert.* **11**, 60–72 (2011)
37. Ohme, R., Matukin, M., Boshoff, C.: Conscious and Subconscious Evaluation of Service Recovery Situation, pp. 65–69 (2016). https://doi.org/10.1007/978-3-319-24148-7_20

Pre-ordering Music Vynils: The Impact of Price and Number of Band Supporters



Sónia Felizardo, Pedro Bem-Haja, Beatriz Casais, Sónia Brito-Costa, and Hugo de Almeida

Abstract The music industry has undergone a significant transformation from physical to digital consumption. This change has altered the sales strategies of products, allowing consumers to purchase them before their market launch. This paper analyzes the effect of pre-order adherence on the digital music platform Bandcamp, basing on the influence of price, number of supporters and their interaction on purchase intention. The sample of this behavioral study was supported by a total of $N = 41$ individuals, with an average age of 26.34 ($SD = 6.279$). Seven-point Likert scales were used to record each participant's level of interest and likelihood to purchase in the face of the stimuli presented to them. For the presentation of forced stimuli, we used 30 layouts of the Bandcamp platform, divided into six conditions = three price levels and two levels of number of supporters to which we wanted to verify their effect on the interest and purchase intention of the experiment subjects. Results show that the number of supporters is a significant predictor of purchase probability and that this probability is higher when participants are in the presence of many supporters. Regarding price, the effect that was expected did not occur, and the low price was not significantly more appealing. Interestingly, the medium price was the least appealing, and the high price was almost as appealing as the low price.

Keywords Price · Supporters · Pre-order · Bandcamp · Consumer behavior

S. Felizardo

CNL-Consumer Neuroscience Laboratory, University of Aveiro, Aveiro, Portugal

P. Bem-Haja

Neuro Laboratory, University of Aveiro, Aveiro, Portugal

B. Casais (✉)

School of Economics and Management, University of Minho, CICS.NOVA.UMinho, Braga, Portugal

e-mail: bcasais@eeg.uminho.pt

S. Brito-Costa

Human Potential Development Center, Polytechnic Institute of Coimbra, Coimbra, Portugal

S. Brito-Costa · H. de Almeida

Research Group in Social and Human Sciences, Coimbra Education School, Polytechnic Institute of Coimbra, Coimbra, Portugal

1 Introduction

Given the technological advances, the emergence of the Internet and, consequently, the new media at the turn of the twentieth to the twenty-first century, the patterns of consumption, production and distribution in the music industry have also changed, leading to the implementation of new methodologies in the process of distribution and availability of products on the market [1]. The most significant advance came with the creation and implementation of the MP3 file in the market, which together with the expansion and progressive development of digital technologies led to the birth of peer-to-peer (P2P) distribution services of free music that later led to the appearance of Web sites for the distribution and sharing of this content [2]. Napster, one of the most famous Web sites, had the advantage of allowing the sharing of numerous audio files, encoded in MP3; however, by operating under the concept of decentralized file sharing, they did not own copyrights for the audio files they were making available [3]. Since the introduction of these data sharing Web sites online, illegally downloaded music (pirated music) has skyrocketed, posing an economic threat to the industry [4]. The music industry has tried to adapt to the Internet era through the creation of online stores such as iTunes, and in recent years by promoting the creation of the first streaming platforms, which, despite the benefits from the consumer's perspective, have problems mainly in terms of sales [5–7]. These difficulties led to the creation of sales methodologies, such as pre-ordering, that would protect the artist by reducing the implicit uncertainty of demand and at the same time benefit consumers through advantages [8, 9]. Thus, the pre-order process became fundamental in the market for its usability in predicting sales after the launch of new products [10] but also for the ease of the consumer to track the order and buy with extra benefits compared to what they would have if they waited until the product was launched in the market [11].

One of the incentives for pre-orders often mentioned among the academic community is the price reduction strategy [12–14]. Some studies also mention the importance of elements such as the details provided in the product information [7, 8, 12, 15] or the innovation presented by the product [11]. Furthermore, the consumer is strongly influenced by culture, subculture, social class, reference groups, family, personality, psychological factors, among others, in addition to cultural trends and the social environment in which he/she is inserted [16]. With the emergence of social networks, one of the most important vectors of influence is the number of followers, which leads to greater perceptions of popularity and greater acceptance [17]. Through the use of the Bandcamp platform, this study intends to verify if the consumer is influenced by the price and by the amount of people who also pre-purchased that album (supporters)—and if the manipulation of these elements results in an increase or decrease in the consumer's interest and probability of purchase.

2 Methods

2.1 Sample

The sample of this behavioral study was supported by a total of 41 participations (21 males and 20 females) from Portugal, between 19 and 54 years old, with an average age of 26.34 ($SD = 6.279$), the most frequent being 24 years old. Regarding academic qualifications, this sample was mostly composed of individuals who had completed the 12th grade (34.1%), a bachelor's degree (29.3%) and a master's degree (24.4%). In what concerns the current professional situation, the sample was essentially represented by individuals who were either employed (48.8%) or students (26.8%). Of the individuals who participated in this study, 90.2% (i.e., 37 participants) were currently single and included participants residing in the districts of Aveiro (14.6%), Braga (9.8%), Castelo Branco (4.9%), Leiria (26.8%), Lisbon (19.5%), Porto (19.5%) and Viseu (4.9%).

2.2 Measures

To assess the opinions and attitudes of the participants in this behavioral study toward the stimuli they were going to be shown, a 7-point Likert scale was used to record each participant's degree of interest and probability of purchase toward the *stimuli* presented to them, in order to verify the impact of these *stimuli* later on the data analysis. Two 7-point scales were also used, one is to measure the participants' interest in the various vinyl records available for pre-order that were shown to them, and another is to measure the probability that the person would buy each of those records. The *Likert* scale has been adapted for this case, serving only to measure the participants' opinion based on *stimuli* and not according to certain statements, since the primary objective of this research is to understand the impact of certain pre-order elements of the Bandcamp platform on consumer behavior, by analyzing different conditions.

2.2.1 Stimuli

The authors used 30 layouts of the Bandcamp platform, divided into six conditions—corresponding to three price levels and two levels of number of supporters—to which the researchers wanted to verify their effect on the interest and purchase intention of the individuals in the experiment. The participants had to evaluate a vinyl LP album in a modified Bandcamp layout, answering the interest this product aroused and the probability of buying it, using as a measure a 7-point scale which was programmed to appear 10 s after being shown the layouts of the Bandcamp platform. These *stimuli* were shown sequentially and randomly, and the participants' response was given



Fig. 1 Example of three layouts designed for the condition few supporters with the conditions low price, medium price and high price (from left to right)

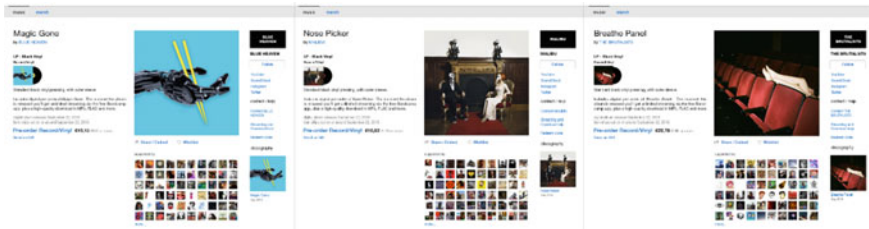


Fig. 2 Example of three layouts designed for the condition many supporters with the conditions low price, medium price and high price (from left to right)

by selecting the number key that best matched their evaluation and probability of buying the product (between a minimum of 1 and a maximum of 7). The structure of the layouts built for implementation in the behavioral study, as *stimuli*, is the same with the exception of the variables on which we intend to measure the impact (price and number of supporters) and the control variables previously selected by the pilot studies. Since each condition was repeated five times, a total of 30 layouts were built.

Figures 1 and 2 show six of the stimuli (or layouts) representing each of the conditions under analysis, for a total of six conditions.

2.2.2 Stimuli Presentation

The presentation of the *stimuli* was done using an experiment builder called OpenSesame [18].

2.3 Design

Each participant was guaranteed to be exposed five times to the same condition (replication principle), where the elements that changed in each repeated condition were only the album covers and the names of albums and bands/artists. In the total experiment, participants were exposed to a total of 30 stimuli. The *stimuli* shown to

the participants followed a random order (randomness principle), so each participant was exposed to the *stimuli* in a different order. Since all participants see all the conditions under study, this is a within-subjects research design, through the association of 3×2 variables, consisting of a total of six conditions.

The purpose of repeating and randomly showing the conditions among the participants was to try to understand if they effectively gave their answer based on the price and the number of supporters that a certain record had in the Bandcamp layout, or if there were other elements that had an impact on that answer such as the variables that were collected in the characterization of the sample and its consumption patterns.

2.4 Procedure

The subjects were participated in the study in an informed and voluntary manner and were informed about the context of the research, the duration of the test and the tasks that would be proposed to them. All participants filled out an informed consent form, in which they authorized that their data would be used in this research. The experiments were divided into three phases: a first phase corresponding to the completion of a questionnaire, a second phase corresponding to the main experiment conducted in OpenSesame, and a third phase in which they were asked to answer a short questionnaire about the main experiment of the research.

In the first phase, participants were asked to answer a total of 13 questions, six of which were related to their socio-demographic data and the remaining seven questions related to cultural and music consumption, specifically. While the participants were answering the questionnaire, the OpenSesame software was being used to prepare the experiment they would perform later, hence the need to meet them in person.

Once the participants completed the questionnaire, they were given a number that was later used to identify the results of their OpenSesame experiment. Before starting the practical experience, the participants were again informed about the tasks they would have to perform through an instruction screen that explained that in the next few minutes they would be viewing a total of 30 static layouts of the Bandcamp platform and that they should observe them as if they were interacting with the platform in a real environment. This screen also explained that after viewing each of the layouts, the participants would have to answer two questions about the previously viewed stimulus and that when they were ready to begin the experiment, to press the start key.

The OpenSesame experiment began with a 1 s fixation point, followed by one of the 30 layouts on the Bandcamp platform, which were presented for a period of 10 s each. After viewing the screen corresponding to the stimulus (layout), the participants were exposed to two successive questions where they evaluated their level of interest in the product shown and the probability of buying the product shown, using a 7-point rating scale. To give the answer they considered most correct in their case, participants pressed the keys from 1 to 7, according to their level of



Fig. 3 Experimental design of the behavioral study

interest in the product and probability of buying it. No answers other than those comprising numbers 1, 2, 3, 4, 5, 6 and 7 were accepted by the program. Each time the participant answered the last of the two questions, and a new loop was started, showing the layouts randomly to all participants. The OpenSesame experiment ended as soon as a new information screen appeared, thanking the participant for their time and inviting them to submit their results (Fig. 3).

To finalize participation in this behavioral study, as soon as participants submitted their results of the experiment in OpenSesame, they were asked to answer a short post-experiment questionnaire where they had, out of the total number of stimuli they were presented with, to indicate how many album covers and album and/or band names they recognized, in a closed-ended type question.

The process of analyzing the results in this section used IBM SPSS Statistics 23 software.

3 Results

A multilevel analysis with the linear model was performed for the dependent variable probability of product purchase. The results showed a main effect on the number of supporters (Wald Chi-square = 10.594, $df = 1$, $p < 0.01$), showing that the number of supporters is a significant predictor of purchase probability and that this probability is higher when participants are in the presence of many supporters. However, there were no price effects (Wald Chi-square = 0.842, $df = 2$, $p = 0.656$), nor intersection between number of supporters and price (Wald Chi-square = 1.399, $df = 2$, $p = 0.497$) (Table 1).

Table 1 Model of the main effect and the intersection between the number of supporters and price on the likelihood of buying a vinyl available on pre-order

Source	Type III		
	Wald Chi-Square	df	Significance
Intercept	271.727	1	0.000
Supporters	10.594	1	0.001
Price	842	2	0.656
Supporters * price	1.399	2	0.497

It can be seen that a medium price is the most likely to lead a participant to have a higher probability of purchase (when the number of supporters is large) or a lower probability of purchase, even compared to a high price (when the number of supporters is smaller) (Fig. 4).

Starting the predictive effect of supporters and crossing with price for purchase probability conditions, the results showed significant differences only between the “few supporters with medium price” condition and the “many supporters with medium price” condition ($p < 0.05$). There was also a decreasing effect on the participants’ probability of purchase between the “few supporters with high price” condition and the “many supporters with high price” condition, but this difference only assumed a statistical trend ($p = 0.064$). In turn, the “few supporters with low price”

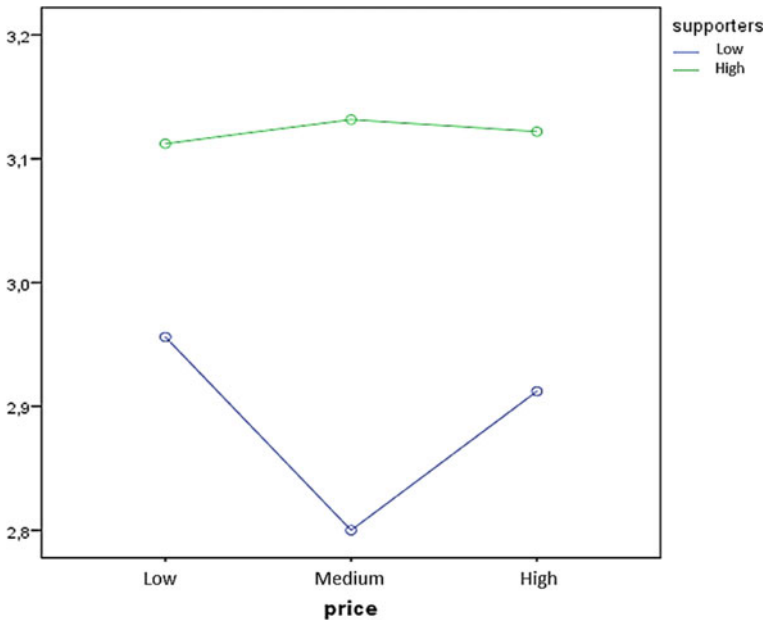


Fig. 4 Graph of the marginal means comparisons estimated based on price and number of supporters for the dependent variable probability of purchase

condition did not obtain a significantly different result from the “many supporters with low price” condition, nor did it assume any statistical trend ($p = 0.154$).

Thus, the overall prediction effect of the supporter’s variable on the probability of buying a physical album (vinyl) is mainly explained by the medium price conditions since for a low price, having many or few supporters has no significant impact on the probability of purchase. Thus, it can be concluded that the condition that combines a low number of supporters with medium price tends to significantly reduce the probability of a product, compared to all conditions that combine a high number of supporters, regardless of price.

Despite not having verified an interaction effect between price and number of supporters, in order to answer the hypothesis that the lower the price of an album on pre-order and the lower the number of supporters, the lower the probability of purchase, it was found that this hypothesis cannot be confirmed with statistical significance. Thus, although a low number of supporters associated with a low price decrease, on average, the probability of buying a vinyl record (compared to all conditions with many supporters), this reduction in the probability of purchase is not significant. Furthermore, a low price in the presence of few supporters increases the probability of purchase, compared to all other conditions that had few supporters (although this increase is not significant). In sum, the condition that guaranteed a greater reduction in the probability of buying a vinyl record was the condition that presented fewest supporters and an average price, unlike the low price, formulated in theory.

4 Discussions and Conclusions

With the avalanche of the digital world leading to a sharp decrease in sales, the music industry has made an effort to reinvent itself [4]. In this sense, the industry has tried to use the online world to its advantage by creating various marketing and sales strategies. One of the best known and most accepted by consumers is the pre-order function, as it allows the industry to manage demand before release while benefiting the consumer with several advantages such as a reduced price. The focus is on creating marketing strategies to optimize pre-order adoption; however, to the best of our knowledge, no study has shown the influence of price, number of supporters and their interaction on purchase intention. The results show that the number of supporters is a significant predictor of purchase probability and that this probability is higher when participants are in the presence of many supporters. These results show us that increasing the number of followers or supporters leads to greater perceptions of popularity and greater acceptance [17]. In fact, in the absence of information that helps consumers perceive the weight of the product based on its attributes and desired utility, consumers tend to focus only on the impressions caused by the product’s image and potential users [19, 20]. This effect is also a manifestation of the old social psychology maxim that people are naturally motivated to build and maintain relationships with others [21]. As far as price is concerned, the expected effect did

not occur, and the low price was not significantly more appealing. Interestingly, the average price was the least appealing, and the high price was almost as appealing as the low price. This result can be explained by the fact that price is also associated with product quality. Thus, we probably have two types of consumers regarding decision making, the “Hagglers” who opt for the affordable and the “Quality connoisseurs who opt for the Good”, the median price being too expensive for the Hagglers and the average quality being too low for the “Quality connoisseurs”. Interestingly, when the median price is accompanied by the existence of many supporters, the effect reported above disappears because the high number of supporters gives the Hagglers confidence to spend a little more and hints to the quality connoisseurs that the quality can be good.

This study not only gives an understanding of purchasing decision making, but also gives a signal to the industry that bands need to work their social networks to increase their supporter base and reward loyalty, because from the results of this study we can see that a large supporter community is a guarantee of a large consumer community. However, more scientific evidence is needed to verify if these data replicate.

References

1. Molteni, L., Ordanini, A.: Consumption patterns, digital technology and music downloading. *Long Range Plan.* **36**(4), 389–406 (2003). [https://doi.org/10.1016/S0024-6301\(03\)00073-6](https://doi.org/10.1016/S0024-6301(03)00073-6)
2. Alexander, P.J.: Peer-to-peer file sharing: The case of the music recording industry. *Rev. Ind. Organ.* **20**(2), 151–161 (2002). <https://doi.org/10.1023/A:1013819218792>
3. Tschmuck, P.: *Creativity and Innovation in the Music Industry*. Springer, Netherlands (2006)
4. Coleman, M.: *Playback: From the Victrola MP3, 100 Years of Music, Machines, and Money*, 2nd edn. Da Capo Press, Cambridge, MA (2005)
5. Aguiar, L.: Let the music play? Free streaming and its effects on digital music consumption. *Inf. Econ. Policy* **41**, 1–14 (2017). <https://doi.org/10.1016/j.infoecopol.2017.06.002>
6. Aguiar, L., Martens, B.: Digital music consumption on the internet: evidence from clickstream data. *Inf. Econ. Policy* **34**, 27–43 (2016). <https://doi.org/10.1016/j.infoecopol.2016.01.003>
7. Bourreau, M., Gensollen, M., Moreau, F., Waelbroeck, P.: Selling less of more? The impact of digitization on record companies Patrick Waelbroeck working paper 12-TS-03. *Selling Less of More? The Impact of Digitization on Record Companies*, pp. 1–24 (2012)
8. Chu, L.Y., Zhang, H.: Optimal preorder strategy with endogenous information control. *Manage. Sci.* **57**(6), 1055–1077 (2011). <https://doi.org/10.1287/mnsc.1110.1335>
9. Loginova, O., Wang, X.H., Zeng, C.: Learning in advance selling with heterogeneous consumers. *Manage. Decis. Econ.* **38**(6), 765–783 (2017). <https://doi.org/10.1002/mde.2815>
10. Moe, W.W., Fader, P.S.: Fast-track: article using advance purchase orders to forecast new product sales. *Mark. Sci.* **21**(3), 347–364 (2002). <https://doi.org/10.1287/mksc.21.3.347.138>
11. Byun, K.-A., Dass, M., Kumar, P., Kim, J.: An examination of innovative consumers’ playfulness on their pre-ordering behavior. *J. Consum. Mark.* **34**(3), 226–240 (2017). <https://doi.org/10.1108/JCM-01-2016-1684>
12. Li, C., Zhang, F.: Advance demand information, price discrimination, and preorder strategies. *Manuf. Serv. Oper. Manage.* **15**(1), 57–71 (2012). <https://doi.org/10.1287/msom.1120.0398>
13. Loginova, O.: Pricing strategies in advance selling: should a retailer offer a pre-order price guarantee? *Rev. Ind. Organ.* **49**(3), 465–489 (2016). <https://doi.org/10.1007/s11151-016-9507-2>

14. Mukherjee, A., Jha, S., Smith, R.J.: Regular price \$299; pre-order price \$199: price promotion for a pre-ordered product and the moderating role of temporal orientation. *J. Retail.* **93**(2), 201–211 (2017). <https://doi.org/10.1016/j.jretai.2016.11.001>
15. Zhao, X., Stecke, K.E.: Pre-orders for new to-be-released products considering consumer loss aversion. *Prod. Oper. Manag.* **19**(2), 198–215 (2010). <https://doi.org/10.1111/j.1937-5956.2009.01092.x>
16. Solomon, M.R., Bamossy, G.J., Askegaard, S., Hogg, M.K.: *Consumer Behaviour: A European Perspective*. Pearson Education Limited (2006)
17. Veirman, M., Cauberghe, V., Hudders, L.: Measuring through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *Int. J. Advert.* **36**(5), 798–828 (2017)
18. Mathôt, S., Schreij, D., Theeuwes, J.: OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* **44**(2), 314–324 (2012). <https://doi.org/10.3758/s13428-011-0168-7>
19. Aguirre-Rodriguez, A., Bosnjak, M., Sirgy, M.J.: Moderators of the self-congruity effect on consumer decision-making: a meta-analysis. *J. Bus. Res.* **65**(8), 1179–1188 (2012). <https://doi.org/10.1016/j.jbusres.2011.07.031>
20. Frattini, F., Bianchi, M., De Massis, A., Sikimic, U.: The role of early adopters in the diffusion of new products: differences between platform and nonplatform innovations. *J. Prod. Innov. Manage.* **31**(3), 466–488 (2014). <https://doi.org/10.1111/jpim.12108>
21. Baumeister, R.F.: Need-to-belong theory. In: Van Lange, P.A.M., Kruglanski, A.W., Higgins, E.T. (eds) *Handbook of Theories of Social Psychology*, pp. 121–140. Sage Publications Ltd. (2012). <https://doi.org/10.4135/9781446249222.n32>

Technologies Applied to Tourism Marketing

360° Technology Applied to Touristic Marketing



Omar Córdor-Herrera and Carlos Ramos-Galarza

Abstract 360° technology for images or videos allows the covering of a 360° vision angle in two axes (X and Y), and it can be seen on a screen in an interactive way. This means that we can see everything that is around the point where the picture or video was taken, up, down, in front, behind and the sides, which allows us to reproduce the stereoscopic human vision. In this article, the usage of this technology is proposed as a tool to enhance the touristic marketing area, giving the users an innovative experience.

Keywords 360° technology · Tourism · Marketing · Virtual reality

1 Introduction

The augmented reality and the virtual reality are different emerging technologies which are being applied in different scopes [1], videos that use a virtual reality device (VR) depict a new means that offer amplified possibilities [2], 360° videos allow spectators to look in whatever direction, while the video is played, these characteristics have allowed that the interest toward the application of the 360° is amplified in a significative way [3]. This article presents a detailed review of 360° technology, operation, devices and applications that can be used, and at the same time, the use of this technology to enhance the fields of tourism marketing.

O. Córdor-Herrera · C. Ramos-Galarza (✉)

Centro de Investigación en Mecatrónica y Sistemas Interactivos MIST/Carrera de Psicología,
Universidad Tecnológica Indoamérica, Av. Machala y Sabanilla, Quito, Ecuador
e-mail: caramos@puce.edu.ec; carlosramos@uti.edu.ec

O. Córdor-Herrera

e-mail: omarcondor@uti.edu.ec

C. Ramos-Galarza

Facultad de Psicología, Pontificia Universidad Católica del Ecuador, Av. 12 de Octubre y Roca,
Quito, Ecuador

2 360° Technology

Current technology allows to repose the way in which we interact with reality, and in that field, 360° technology in images and videos is the most representative because it allows to cover a vision angle of 360° in two axles (X, Y) when an image is captured which when visualized on a screen, the user is given a creative experience in relation to the image and audio, creating a new audiovisual language.

The advantage of using 360° technology is that it allows viewers to view the productions of a tourist place several times, allowing them to distinguish different points of view of the place [4] every time, this technology is being applied, to the business and institutional fields [5], for example, at the same time than the 360° application, the so called 360° routes have also been created, which are the digital solutions that allow to create experiences such as virtual interactive tours of properties or places of interest, allowing at the same time to explore profoundly the visual environments in an interactive, simple and effective way [6]. The routes are based on images that turn into panoramic views, allowing the complete vision of a space where the user goes through the real environment sensation through the control of movement in the area, with the use of a pointer or computer keys [7].

2.1 *Devices for the Creation of 360° Proposals in Tourism*

In order to create 360° photographs and videos, it is necessary to count on devices that are adequate for this type of job such as the 360 camera, which is a device that can make pictures and videos thanks to two wide angle lenses, placed one behind the other. Through this, when the shot is made with the 360 camera, it can capture with one lens the frontal part and with the other lens the rear part, because of them being wide angle lenses not only do they capture a front and rear image but also the floor, roof and lateral views (Fig. 1).

Once the picture is taken through the lenses of the 360 camera, an image file will be obtained with the two round pictures: the front and the rear [8] (Fig. 2). In this moment, a process where the front and rear pictures are combined will occur in order to obtain one image, and this process is called stitching which literally makes reference to “stitching” [8].

There currently are different cameras that allow to obtain takes in 360, and in Table 1, different 360° cameras with their characteristics will be seen [9].

Panoramic View by 2 Spherical Lens



Fig. 1 Zetronix 360 camera



Fig. 2 Stitching process in a 360° process

2.2 Virtual Reality Glasses and 360° Technology in Tourism

The virtual reality glasses allow an immersive experience where the user can have a 360° access to touristic destinations where there are access difficulties or simply for the user to have a first experience before choosing a touristic destination (Fig. 3).

Table 1 Cameras 360°

Camera	Characteristics
SOONPAM 360 mobile camera	2048 × 1024 for images 1080P for videos Extremely portable from 1.6 × 1.6 inches Allows Android 5.0 with micro USB and also type C connector Allows 360 transmission in Facebook/Twitter and YouTube
Samsung gear 360	Lenses F2.0 capture 4K videos or 30 MP pictures Resists occasional splatters and the tempered glass protects from scratches to the lens Plays, cuts and shares videos directly on a Galaxy device or compatible PC Micro SD card (sold separately) required to work
Zetronix 360 camera	960P @ 30fps for up to 2 h 3008 × 1504 JPG image resolution Built-in Wi-Fi for wireless transfer Accepts micro SD (TF) cards up to 32 Camera, mounting support and tripod included in package
FSTgo VR 360	Fisheye lens of 230° or a mirror, in different directions, different shots stitched among themselves with a panoramic view of 360 degrees 4 K 30fps 1920 * 1440 @ 30fps effective pixels Ultra HD Panorama 3D VR Camera 1200mAh rechargeable Li-ion battery 20 M in all-direction through Splashproof remote control to WIFI APPLICATION: 802.11 b / g
Gear Pro360	Built-in microphone. Camera records 1080p HD videos with high fidelity audio Resolutions can be changed easily: 12MP, 8MP, 5MP or 3MP LCD screen of 1.5 Includes rechargeable battery and USB cable to charge and transfer data Includes Hype360° camera kit, waterproof protection, handlebar mount, universal base mount and mounting adjustable straps. Put it in your helmet, bicycle or surfing board!
Andoer 360	Slow movements to help it capture a special moment Supports TF SD cards of up to 128 GB (not included) Comes with an airtight case, 30 M maximum Video resolution: 2448P(30fps); 1440P(60fps) Image resolution: 16 M(4096 * 4096) Includes bicycle mount, USB cable, helmet base and waterproof cover
Zision 360 4K	Virtual reality camera with 4K resolution 1050 mA, battery, working time up to 90 min Waterproof until 30 m Equipped with 170° wide angle lens Video resolution up to 4K 1080P 30 fps Stabilizing function and built-in microphone Smartphone remote control with powerful application

Fig. 3 Representation of the use of virtual reality glasses



The virtual reality glasses allow the user to know details of tourist sites before visiting them, this feature is of great importance during the 2019 and 2020 pandemic, in addition to other circumstances in which it is impossible for the human being to access the site because this device offers a very real experience of the different touristic sites that can be visited throughout the world [10].

2.3 Applications to Create Touristic Proposals with 360° Technology

There are different applications that allow to develop technological 360° proposals; for example, the Google Street View application allows to create total representations of the diverse spaces in which the user is located in and through the use of the smartphone camera recreates the touristic places of interest [11].

Another application that allows the creation of circular images and videos of places of interest in the touristic context is DMD Panorama. This application works properly with low illumination, and its interface is easy to use for the user. Finally, Matterport is a dynamic application and of easy use, and it allows to do 360° scenery routes of touristic places that can have interesting designs for the user [12].

2.4 Tourism Marketing and Technology 360°

Tourism marketing is the promotion of places carried out by companies in the hotel sectors, leisure venues, restaurants, transport companies and other public bodies [13] for which tourist advertising is used, which is considered an activity that generates influence and strives to persuade people to travel to a certain place, proposing reasons that awaken them the desire to know it [14]; however, the 2020 tourist is different from

years ago because the new consumer likes personalized information, investigates in Google, consults social networks, looks for videos and valuable information about the different places you would like to visit.

It is here where virtual reality and 360° technology take center stage since with the utility of these technologies, tourists can experience the places they want to visit before embarking on the real trip, which leads to better customer service [13] and allows the promotion of tourist places in a different way, directly benefiting tourist marketing, as is the case of tourist promotion in the city of Tunja through the platform layer which allows applying a variety of tourist attractions to points of interest actions that provide the user with more information about tourist places in an interactive way [15, 16].

Summarizing, new technologies have changed the behavior of tourists and have revolutionized the information search process, which is directly related to decision-making, since the more information the user of the place to visit has, the easier and faster their choice will be [17].

3 Conclusions

360° technology is a great tool for the creation of touristic resources, because it allows the user to have a first encounter with the details of the touristic sites that are of the user's interest.

In this article, some elements of the benefits of this type of technology have been shown, such as its use in times when the user cannot travel, like in the case of the recent pandemic, the diverse applications that can be used to create touristic 360° content and the devices used in this field.

Tourists are currently looking for detailed information about the possible places to visit, the more information they have about the place the easier the decision is, so for the use of 360° technology for tourism marketing, it has great potential because it allows the user to virtually visualize places that they would like to visit before seeing them in real life. In following investigations, the creation of 360° content of touristic Latin America sites is of interest, as is for them to be applied in investigation processes in which the experience of the user with the created resources can be analyzed.

References

1. Condor-Herrera, O., Acosta-Rodas, P., Ramos-Galarza, C.: Augmented reality teaching resources and its implementation in the teaching-learning process. *Adv. Human Factors Train. Educ. Learn. Sci. AHFE* **269**, 149–154 (2021)
2. Arents, V., de Groot, P., van Stralen, K., Struben, V.: Use of 360° virtual reality video in medical obstetrical education: a quasi-experimental design. *BMC Med. Educ.* **21**(1), (2021)
3. Snelson, C., Hsu, C.Y.: Educational 360-degree videos in virtual reality: a scoping review of the emerging research. *TechTrends* **64**(3), 404–412 (2020)

4. SIMPLEMENTE [Simply] July 2nd, 2021. [Online]. Available: <https://www.simplemente.net/tecnologia-360-simplemente/>
5. de la casa, J.M.H., Caerols, M., Sidorenko, P.: Virtual reality and 360°video in business and institutional communication **18**(2), 177–199
6. Gomez, J., Gonzalez, D., Gonzalo, A., Herrero, A., Muñoz, Á., Del Val Riaño, J.: Reconstrucción 3D y Realidad Virtual en Criminología [3D Reconstruction and Virtual Reality In Criminology]. Universidad Salamanca Editions (2010)
7. Vargas, D., Otero, J.: Desarrollo e Implementación de Recorridos 360° en portales Joomla!, [360 routes implementation and development, Joomla! portals] *Sci. Technica* **20**(1), 61–69 (2015)
8. Audiovisual.com Group: CÁMARA 360° ¿Qué es y cómo funciona una CÁMARA 360 grados? Foto y Vídeo inmersivo, (360 CAMERA, what is it, how does a 360 degree CAMERA work?) (2021)
9. VEER VR BLOG: VEER VR BLOG, July, 16th 2018. [Online]. Available: <https://veer.tv/blog/es/top-10-camaras-360-que-puedes-conseguir-por-menos-de-100/>. Last accessed 2 July 2021
10. Schiopu, A., Hornoui, R., Padurean, A., Nica, A.: Virus tinged? Exploring the facets of virtual reality use in tourism as a result of the COVID-19 pandemic. *Telematics Inform.* **60**, 101575 (2021)
11. Salgado, A., Li, W., Caridi, I., González, M.: Street context of various demographic groups in their daily mobility. *Appl. Netw. Sci.* **6**(1), 43 (2021)
12. Vasquez, C., Tan, N., Sadalgi, S.: Home studio: a mixed reality staging tool for interior design. In: Conference on Human Factors in Computing Systems Proceedings (2021)
13. Galiana, P.: Innovation and Entrepreneurship Business School. [En línea]. Available: <https://www.iebschool.com/blog/marketing-digital-marketing-digital-2-marketing-digital/> (2021)
14. Aires, G., Nicolau, K.: Marketing turístico internacional. *La Marca Brasil. Estudios Perspectivas Turismo* **19**(2), 241–267
15. Callejas, M., Quiroga, J., Alarcón, A.: Ambiente interactivo para visualizar sitios turísticos, mediante realidad. *Ciencia Ingeniería Neogranadina* **21**(2), 91–105 (2011)
16. Amaya, L., Santamaria, L., Mendoza, J.: Virtual environment for Android platforms related to historic places of the city of Tunja (Colombia). *Revista Virtual Univ Católica Norte* **48**, 276–294 (2016)
17. Lemos, E., Moscardi, E., Alves, M., Nakatani, M.: The relationship between the tourist information and the information and communication technologies: an analysis of the Latin American tourism literature. *Estudios Perspectivas Turismo* **27**(3), 569–587 (2018)

Digital Marketing and User-Generated Content: A Case Study of Vidago Palace Hotel



Irina Clara, Teresa Paiva , and Elisabete Paulo Morais 

Abstract Digital marketing is essential for hotels. It is important to have a social media strategy and, simultaneously, to be aware of the guests' opinions and reviews, be conscious of their satisfaction. In this sense, the content, and interactions of the different social media of Vidago Palace Hotel were analysed, and some suggestions for improvement were given. On the other hand, the hotel reviews in the main online review platforms were analysed, to understand the level of guest satisfaction. The results show that there should be more care at the social media strategy level, posting with more frequency. However, the reviews are excellent, which shows that the quality of service provided by the hotel. The exploratory nature of the study is one of its limitations, added by the fact that is based on a specific date that analysed its last 10 publications.

Keywords Digital marketing · Social media · Online review platforms · User-generated content · Hotel

1 Introduction

Since the emergence of Web 2.0 platforms, there has been more active participation of tourists in social media through ratings and comments of their trips [1].

The use of social media has become increasingly relevant as part of the tourism experience, as it has changed the way information about travel and tourism is published and shared [2]. Social media are the most appropriate tools for tourism

I. Clara
Polytechnic Institute of Bragança, Bragança, Portugal

T. Paiva
Technological and Management School, NECE Research Center in Business Sciences, Guarda Polytechnic Institute, CI&DEI, Guarda, Portugal

E. P. Morais (✉)
UNIAG (Applied Management Research Unit), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
e-mail: beta@ipb.pt

companies to communicate and build long-term relationships with their consumers, allowing them to draw relevant knowledge and information from the market [3]. So, it is very important to have a well-defined and consistent social media strategy, to give all the necessary information to consumers and keep them interested in the published content.

Online ratings are the most important form of electronic word-of-mouth (eWOM) and have an important effect on consumer attitudes and behavior. These reviews positively impact hotel occupancy and hotel revenue, leading to better performance [4]. Thus, it is necessary to monitor guests' opinions, understand if the service provided meets the hotel's positioning, and create a long-term relationship with guests.

2 Literature Review

2.1 Technologies and Internet in Tourism

Information and Communication Technologies (ICT) increasingly play a key role in the competitiveness of tourism organisations and destinations and the industry as a whole [5]. Technologies have conceptualized the way travel is planned [6], and services and experiences are created and consumed [7]. The rise of the Internet has led to the emergence of several opportunities for improvement at communication, distribution channels, and transactions, enabling more effective segmentation and positioning strategies [8]. With the Internet, consumers learned new ways of interaction and consumption [8], and the consumer has become a producer, or co-creator, of tourist information, images and experiences [9].

Thus, technologies are transforming management and marketing in the tourism sector from a static and utilitarian sense, in which managers and tourists used technology as a tool, to a conceptualization in which tourism markets and actors "shape and are shaped by technology" [10].

2.2 Digital Marketing in Tourism

Digital marketing develops in parallel with the advancement of technologies [11]. According to American marketing association (AMA), digital marketing is the use of digital channels to promote a brand or to reach consumers and can be executed on the Internet, search engines, social media, or mobile devices [12]. Digital marketing has several advantages. It allows reaching a larger audience, with greater speed [13], without geographical and temporal constraints [14]. Another advantage of digital marketing is quantification, as it allows greater flexibility to collect and measure results than traditional marketing [13, 14]. Also, it allows a greater involvement, with more participation of the consumer, and the possibility of reaching new consumers

through personalization. Also noteworthy is the network effect, which reflects the need to monitor interactions between users since this effect can lead to both the construction of a company and its destruction [14–16].

Digital marketing has created new ways to get relevant information about organisations, and hotels [17]. Proper use of digital marketing tools allows hotel managers to be aware of and track consumers' attitudes, opinions, and satisfaction, allowing them to provide better quality services. It is important that hotels make strategic use of digital marketing tools and use the information they offer to improve their performance [16].

2.3 *User-Generated Content*

Web 2.0, also called the social age, consists of a dynamic, interactive, and participatory consumer-focused Internet whose content is generated by the users themselves [1]. Evaluations of other travellers in forums, blogs, websites, social media, and video-sharing sites for the views of other online users are called user-generated content (UGC) platforms [18]. So, UGC refers to any digital content that is voluntarily created and published, including videos, images, and user posts on different social media [19]. The biggest advantage of UGC is its immediacy [8].

UGC platforms provide personal information about a particular destination. These websites allow potential consumers to exchange information, leave opinions, recommendations and suggestions about destinations, tourism products and services, often with ratings of a particular organisation [8, 20]. Tourists' comments and ratings function both as a source of information and as a recommendation method, as eWOM [21].

Websites such as YouTube, TripAdvisor, and Expedia allow users to share and rate their travel experiences for free, so that each person who comments about a destination/product becomes an integral part of the marketing process [22]. The emergence of these blogs and their development have enabled a more direct engagement of consumers with suppliers, increasing their (consumers') experience and trust in using the internet [8]. The information exchanged and the proximity of consumers in their interactions leads to organisations losing control of what is being said about them, as consumers can influence other potential customers [20].

Currently, consumers trust website reviews (UGC) more than advertising done by the organisations, as these sources are considered genuine and not business-focused [1, 8, 20], with blogs being seen as a more credible source of information than traditional marketing communication [23–26].

These blogs can be used for consumer profiling, acquisition and interaction with consumers, brand reinforcement and awareness, and reputation management. Also, it allows for real-time adaptations by tourism actors, allowing them to broaden their knowledge of the market and competition, and even serves as a tool for internal communication and knowledge management within tourism organisations [8, 21].

2.4 Social Media and UGC Influence on Tourist Behavior

User-generated content has a significant impact on decision making [27]. Through participation in social networks and platforms, consumers generate value and services to each other [28]. Thus, UGC and eWOM have registered an increasing influence in the different phases of travel, from information search to purchase [29, 30]. Tourists contribute to the construction of a destination's image of other tourists by sharing their experiences [25]

Interpersonal influences and eWOM are ranked as the most relevant sources of information in consumers' decision making process. Valence (descriptions or numerical ratings) and volume (number of reviews) are the factors of online reviews that most influence the decision of which destination to travel to [26]. Many tourists base their travel plans on the ratings and recommendations of other users and the shares made on social media about the destination [22]. Social media causes a rapid expansion of eWOM, and this information and opinions, whether positive or negative, influence tourists' behavior when making decisions about their travel [26].

3 Methodology

To understand the digital marketing strategy used by Vidago Palace Hotel, from the digital platforms consumer's point of view, quantitative and qualitative methods were used.

Initially, the hotel's website was analysed, to identify its strengths and weaknesses.

Subsequently, the social media (Facebook, Instagram, Twitter, YouTube, and Pinterest) and the hotel's ratings on online review platforms (Booking, TripAdvisor, and Google) were examined.

3.1 Vidago Palace Hotel

Vidago Palace Hotel (VPH) is a hotel establishment located in the Vidago Natural Park, belonging to the municipality of Chaves, district of Vila Real, Portugal. It was first opened in 1910, having closed in 2006 and reopened in 2010, one hundred years after its inauguration.

The VPH is a 5-star hotel that holds the "Leading Hotels of the World" distinction, and has an offer composed of 70 rooms and suites, library, indoor and outdoor pools, 4 restaurants and 4 bars (gastronomy), thermal spa and golf course, and is one of the most emblematic tourist destinations in the Northern Region of Portugal [31, 32].

3.2 *Organic Search Results*

By searching for the keywords “Hotel in Chaves”, the results mainly show booking platforms. In Booking.com and Trivago, the hotel does not appear. In Destinia, the hotel appears at the bottom of the first page, in “accommodations nearby”. This is because hotel’s location is defined at Vidago and not Chaves. So, it is suggested that one of the keywords in the hotel description include “Chaves”. There is also a list of 56 hotels given by Google itself, in which the VPH appears in last place on the first page.

Through the keywords “Hotel in Vidago”, there is also a strong presence of booking platforms. In Booking.com, VPH appears in second place, but it appears first in Trivago and Google’s list. The official website and Facebook of the hotel appear on the first page of the results.

Searching for “Vidago Palace Hotel”, the first two results are ads, from booking platforms (Booking.com and Trivago, respectively). The first organic result is the VPH website, followed by booking platforms and the hotel’s Facebook page. It is noteworthy that one of the results refers to the website “The Leading Hotels of the World”, a distinction that the hotel holds.

3.3 *Website*

The quality of a website depends on three main aspects: usability, graphic design, and navigability. Usability includes content (quantity and quality of information), ergonomics (flexibility and speed), processes, errors, adaptation (responsiveness), interactivity (contacts, newsletter, and comments) and distribution/marketing (e-commerce function: online booking possibility). Graphic design comprises aspects such as aesthetics, brand image and clarity, while navigability concerns the navigation menu or system [33].

Considering these aspects, Table 1 shows the main positive and negative aspects of the VPH website.

3.4 *Social Media*

The last 10 contents published on Facebook, Instagram, Twitter, and YouTube were analysed. A Pinterest albums summary was made, identifying followers and content. Here, it is not possible to know the publication and album creation dates. This analysis was performed on 30 June 2021.

About Facebook (<https://www.facebook.com/Vidago.Palace.hotel.Spa>), VPH makes an average of 0.45 posts per day (Fig. 1). The most liked posts are those that show the exterior of the hotel, including promotional videos. Of the last 10

Table 1 VPH website evaluation

Strengths	Weaknesses
Information related to COVID-19 (top bar)	Home page: little information
Horizontal menu: categorised information	Horizontal menu: organisation (the “Restaurants and Bars” page appears first than the “Hotel” page)
Content (quantity and quality of information)	No account is taken of F-Pattern (layout)
Speed and flexibility (ergonomics)	The “Hotel” page presents the hotel and its history, but has no mission, vision, or values of the hotel
Search button	Visual appearance of the website is not aligned with the positioning of the hotel (luxury, excellence): layout, colours
Possibility of online booking	Slide show gallery (stacking would be more appropriate)
Responsiveness	The clean and safe icon refers to the general page of Turismo de Portugal, not containing the hotel’s specific security measures (available on the “COVID-19 protocol” page)
Link to social media	
Contact page and newsletter (interactivity)	
Information request	
Written and explicit location	
Information about experiences in the region	

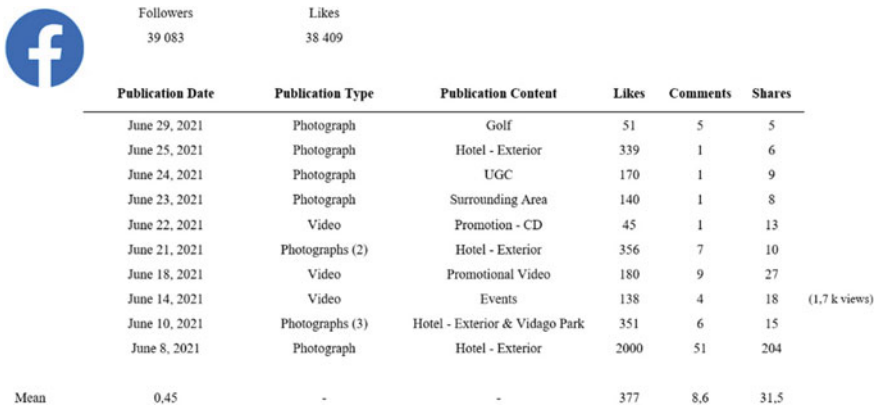


Fig. 1 VPH Facebook publications

publications, 5 show the exterior of the hotel. So, it does not seem to exist a social media management strategy. The hotel could manage it by posting a greater variety of content, for example, the hotel (exterior and interior), the Vidago Park, the restaurants and bars, UGC content, golf and the surroundings (experiences in the region).

This social network has about 39 thousand likes and followers. However, about the video published on June 14, only a small percentage of the followers have seen the video, and few have posted likes or left comments. The last 10 posts, show an average of 377 likes, 9 comments, and 32 shares. So, the hotel should try to ask questions in the descriptions of their posts to encourage their followers to interact. Note that all descriptions of publications are placed in Portuguese and English, and

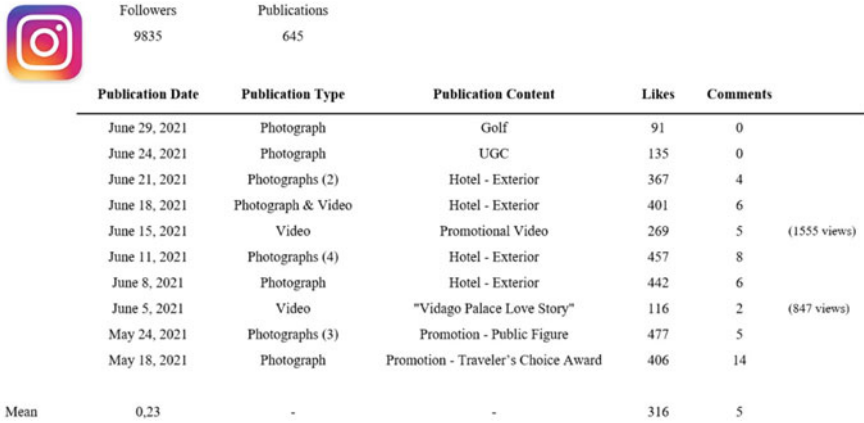


Fig. 2 VPH Instagram publications

the VPH likes all comments left on its publications, responding whenever there are doubts or questions.

On Instagram (<https://www.instagram.com/vidagopalacehotel/>), VPH has about 10 thousand followers (Fig. 2), significantly fewer than on Facebook. However, the number of interactions is similar, and most publications have more likes on Instagram than on Facebook. On Instagram, posts have an average of 316 likes and 5 comments. However, the average number of daily posts is 0.23, which means that this social network is updated every 4 days. Noteworthy is a photo taken by a tourist (UGC), published on the hotel's Instagram.

On both Facebook and Instagram, the hotel uses hashtags frequently, but it uses the same hashtags in most publications. However, since hashtags allow for categorizing information according to a topic or theme while simultaneously functioning as keywords [34], they could be used to categorize different content.

On YouTube (<https://www.youtube.com/vidagopalacehotel/>), VPH has 405 subscribers, and most of the videos published have between one and two thousand views. However, the last activity is from November 2019. In 2016, the hotel had a good frequency of videos, but, currently, it makes no sense to have this social network mentioned on the hotel's official website, as it is outdated.

On Pinterest (<https://www.pinterest.pt/vidagopalace/>), the hotel has 40 followers and 83 monthly viewers. However, it has only 3 albums created to which it has added few pins (images). So, to keep this social network active, it should create albums, for example, for the restaurants and bars, for the spa, for golf, and for events.

About Twitter (<https://twitter.com/VidagoPalace/>), VPH has 439 followers. However, the social network has not been updated since December 2020, and most of the posts are links to Instagram, and most publications have no likes, comments, or shares.

3.5 Online Review Platforms

About online review platforms, Booking, Expedia, TripAdvisor, Trivago, and Google reviews on VPH were summarised. Other platforms such as Destinia were not included because the reviews they provide are based on TripAdvisor reviews.

The evaluations of the VPH are excellent in all platforms analysed (Fig. 3).

Although several elements are well evaluated, the cleanliness stands out, a key point in times of pandemic. Also noteworthy is the staff/service, also with the highest scores. We conclude, then that guests are satisfied with the services provided and with their experience at Vidago Palace Hotel, evaluating it with high scores and recommending it to other tourists. As we know, these comments made by other guests are seen by tourists as more credible than the communication made by the company [23–26]. Tourists consult these endorsements as a source of primary information that reduces the risk and uncertainty of online purchase processes for travel and accommodation-related products [25, 35].

Tourists who consult these platforms will have one more reason to choose VPH, and these high ratings can influence their decision in a positive way.

The hotel should keep an eye on these reviews and try to respond to the positive comments to strengthen its relationship with the guests. On the other side, it should resolve any complaints or negative comments not to lose those consumers, but rather create a friendly relationship with them.

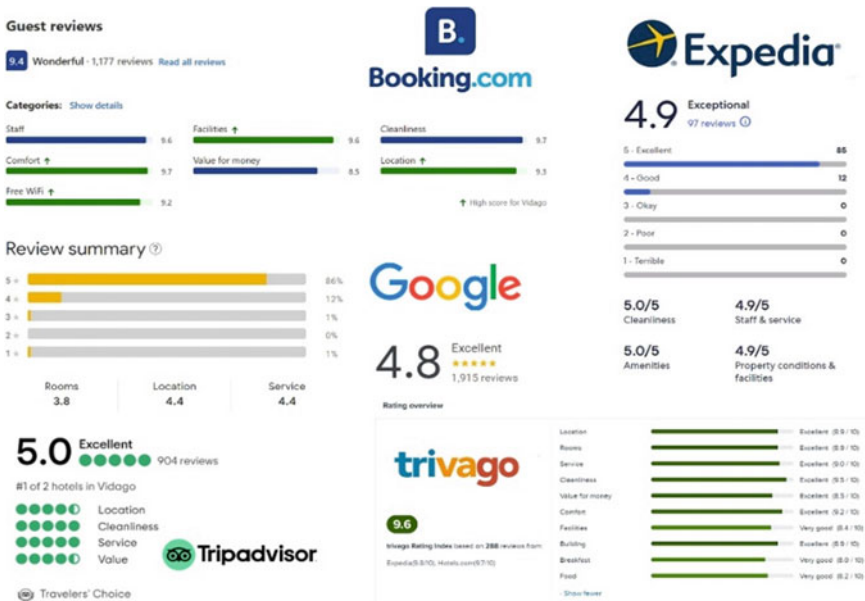


Fig. 3 Online reviews of VPH

4 Conclusion

Posts made on social media and online reviews made by other tourists impact the guests' purchasing decisions. Thus, care should be taken to keep social networks active and consistent with the content posted. Likewise, there should be constant monitoring of the online tourists' reviews since they are an important source of information for hotels.

Vidago Palace Hotel makes, on average, a publication every 2 days on Facebook. In comparison, on Instagram it publishes every 4 days, approximately. The hotel should publish content 5 times a week on Facebook, and 3 times a day on Instagram [36]. Thus, it is suggested that the hotel, on Instagram, try to make at least one publication in the feed, and 2 stories per day. Also, it should pay attention to the content published, show several places, and give different information. The hotel should also be careful to use specific hashtags for each content to categorise the different photos/images. Regarding the other social networks (Twitter, YouTube, and Pinterest), since they are not frequently updated, they should not be promoted on the hotel's website.

Regarding the online review platforms, the ratings are excellent, highlighting the cleanliness and the staff. So, the hotel should follow these comments and try to respond to strengthen the relationship with the guests so that they want to come back.

Acknowledgements UNIAG, R&D unit funded by the FCT—Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education. UIDB/04752/2020.

References

1. Antonio, N., Correia, M.B., Ribeiro, F.P.: Exploring user-generated content for improving destination knowledge: the case of two world heritage cities. *Sustain* **12**, 1–19 (2020). <https://doi.org/10.3390/su12229654>
2. Garrido-Moreno, A., García-Morales, V.J., Lockett, N., King, S.: The missing link: creating value with social media use in hotels. *Int. J. Hosp. Manage.* **75**, 94–104 (2018). <https://doi.org/10.1016/j.ijhm.2018.03.008>
3. Escobar-Rodríguez, T., Carvajal-Trujillo, E.: An evaluation of Spanish hotel websites: informational vs. relational strategies. *Int. J. Hosp. Manage.* **33**, 228–239 (2013). <https://doi.org/10.1016/j.ijhm.2012.08.008>
4. De Pelsmacker, P., van Tilburg, S., Holthof, C.: Digital marketing strategies, online reviews and hotel performance. *Int. J. Hosp. Manage.* **72**, 47–55 (2018). <https://doi.org/10.1016/j.ijhm.2018.01.003>
5. Buhalis, D., Law, R.: Progress in information technology and tourism management. 20 years on and 10 years after the internet. The state of eTourism research. *Tour Manage.* **29**, 609–623 (2008)
6. Neuhofer, B., Buhalis, D., Ladkin, A.: A typology of technology-enhanced tourism experiences. *Int. J. Tour Res.* **16**, 340–350 (2014)

7. Stamboulis, Y., Skayannis, P.: Innovation strategies and technology for experience-based tourism. *Tour Manage.* **24**, 35–43 (2003)
8. Akehurst, G.: User generated content: The use of blogs for tourism organisations and tourism consumers. *Serv. Bus.* **3**, 51–61 (2008). <https://doi.org/10.1007/s11628-008-0054-2>
9. van der Zee, E., Bertocchi, D., Vanneste, D.: Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of TripAdvisor data as support for destination management. *Curr. Issues Tour* **23**, 175–196 (2020). <https://doi.org/10.1080/13683500.2018.1491955>
10. Sigala, M.: New technologies in tourism: from multi-disciplinary to anti-disciplinary advances and trajectories. *Tour Manage. Perspect.* **25**, 151–155 (2018). <https://doi.org/10.1016/j.tmp.2017.12.003>
11. Wirawan, F.A., Oktivera, E.: Analysis on the implementation of digital marketing towards motorbike transport service case study: GO-JEK (online taxi motorbike) Jakarta, Indonesia (2015)
12. American Marketing Association: Digital marketing archives. In: *Digit. Mark.* <https://www.ama.org/topics/digital-marketing/> (2021)
13. Parreira, C.G.: Can't skip Portugal: a influência do marketing digital na escolha de um destino turístico (2018)
14. Lousada, V.I.M.: O papel das redes sociais para a promoção do Hotel Anantara Vilamoura & Algarve Resort (2021)
15. Chaffey, D., Smith, P.R.: *Emarketing Excellence: Planning and Optimizing Your Digital Marketing* (2013)
16. Tarazona-Montoya, R., Peris-Ortiz, M., Devece, C.: The value of cluster association for digital marketing in tourism regional development. *Sustainability* **12**, 9887 (2020). <https://doi.org/10.3390/su12239887>
17. Smithson, S., Devece, C.A., Lapiedra, R.: Online visibility as a source of competitive advantage for small- and medium-sized tourism accommodation enterprises. *Serv. Ind. J.* **31**, 1573–1587 (2011). <https://doi.org/10.1080/02642069.2010.485640>
18. Lam, J.M.S., Ismail, H., Lee, S.: From desktop to destination: user-generated content platforms, co-created online experiences, destination image and satisfaction. *J. Destin. Mark. Manage.* **18**, 100490 (2020). <https://doi.org/10.1016/j.jdmm.2020.100490>
19. Rajamma, R.K., Paswan, A., Spears, N.: User-generated content (UGC) misclassification and its effects. *J. Consum. Mark.* **37**, 125–138 (2020). <https://doi.org/10.1108/JCM-08-2018-2819>
20. Hur, K., Kim, T.T., Karatepe, O.M., Lee, G.: An exploration of the factors influencing social media continuance usage and information sharing intentions among Korean travellers. *Tour Manage.* **63**, 170–178 (2017). <https://doi.org/10.1016/j.tourman.2017.06.013>
21. Da Luz Cerdeira, J.H., Fernandes, A.: Study of the image of the tourist region of Lisbon | Estudo da imagem da região turística de Lisboa. *RISTI Rev. Iber. Sist. Tecnol. Inf.* **2020**, 89–111 (2020). <https://doi.org/10.17013/risti.40.89-111>
22. Bounab, Y., Oussalah, M., Beddiar, D.R.: Impact of quality of images on users behavior on social media. In: *2020 10th International Conference on Image Processing Theory, Tools and Applications, IPTA 2020* (2020)
23. Bigne, E., Fuentes-Medina, M.L., Morini-Marrero, S.: Memorable tourist experiences versus ordinary tourist experiences analysed through user-generated content. *J. Hosp. Tour Manage.* **45**, 309–318 (2020). <https://doi.org/10.1016/j.jhtm.2020.08.019>
24. Hoffman, E., Daugherty, T.: Is a picture always worth a thousand words? Attention to structural elements of eWOM for consumer brands within social media. *ACR North Am. Adv.* **41**, (2013)
25. Lozano-Monterrubio, N., Huertas, A.: The image of Barcelona in online travel reviews during 2017 Catalan independence process. *Commun. Soc.* **33**, 33–49 (2020). <https://doi.org/10.15581/003.33.3.33-49>
26. Martín-Fuentes, E., Nieto Ferrando, J., Marine-Roig, E., Ferrer-Rosell, B.: From blockbuster to neighbourhood buster: the effect of films on Barcelona. *Sustainability* **12**, 2290 (2020). <https://doi.org/10.3390/su12062290>

27. World Bank Group: *The Voice of Travelers : Leveraging User-Generated Content for Tourism Development*. Washington DC (2018)
28. Sigala, M.: E-service quality and Web 2.0: expanding quality models to include customer participation and inter-customer support. *Serv. Ind. J.* **29**, 1341–1358 (2009)
29. Filieri, R., Acikgoz, F., Ndou, V., Dwivedi, Y.: Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention. *Int. J. Contemp. Hosp. Manage.* **33**, 199–223 (2020). <https://doi.org/10.1108/IJCHM-05-2020-0402>
30. Iglesias-Sánchez, P., Marisol, C., Carmen, J., Carlos, D.L.H.-P.: Instagram as a co-creation space for tourist destination image-building: Algarve and Costa del Sol case studies. *Sustainability* **12**, 2793 (2020). <https://doi.org/10.3390/su12072793>
31. Vidago Palace: *Historia Santiago*. <https://www.vidagopalace.com/pt/hoteis/historia/> (2021)
32. Vidago Palace: *Localização*. <https://www.vidagopalace.com/pt/hoteis/localizacao/> (2021). Accessed 15 Jun 2021
33. Martínez-Sala, A.-M., Monserrat-Gauchí, J., Alemany-Martínez, D.: User usable experience: a three-dimensional approach on usability in tourism websites and a model for its evaluation. *Tourism Manage. Perspect.* **33**:100579 (2020). <https://doi.org/10.1016/j.tmp.2019.100579>
34. Gon, M.: Local experiences on Instagram: social media data as source of evidence for experience design. *J. Destin. Mark. Manage.* **19**, (2021). <https://doi.org/10.1016/j.jdmm.2020.100435>
35. Buhalis, D., Parra López, E., Martínez-González, J.A.: Influence of young consumers' external and internal variables on their e-loyalty to tourism sites. *J. Destin. Mark. Manage.* **15**, 100409 (2020). <https://doi.org/10.1016/j.jdmm.2020.100409>
36. Marques, V.: *Marketing Digital de A a Z* (2021)

Websites Usability Evaluation of the Terras De Trás-Os-Montes Hotels



Elisabete Paulo Morais , Carlos Rompante Cunha , and Arlindo Santos 

Abstract Website is a bridge between users and online information. It is extremely important in terms of marketing and must be designed according to the rules of usability, especially in hotel industry. Websites with high usability value will be accessed by more users. Therefore, building a useful website is important. This study aims to evaluate, from the point of view of usability, the websites of hotel establishments in Terras de Trás-os-Montes, a region located in the north of Portugal.

Keywords Hotel · Website · Usability · TTM

1 Introduction

In today's competitive world, a website can be considered a fundamental aspect of an organization's competitiveness. In addition to visual aesthetics, the usability of a website is a strong determinant of user satisfaction and pleasure.

The tourism sector has undergone many changes in marketing and procedures due to the widespread use of the web and its tools. In this new reality, hotel websites represent important tools, namely as an exclusive communication channel [1].

In a context of growth in the volume of revenue generated by the tourism sector in the national economy in recent years and in the current context of the Covid-19 pandemic, the analysis of the online presence of hotels, as economic agents in this sector, appears to be a real need. This presence, mostly embodied in the use of the website, can be scientifically analyzed using instruments that integrate different perspectives, with usability being one of the most important.

E. P. Morais (✉) · C. R. Cunha
UNIAG (Applied Management Research Unit), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
e-mail: beta@ipb.pt

A. Santos
Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

Usability is a process of matching user needs to enable the use of a product or system. There are many methods that can be used to evaluate a website. Usability is one of the main evaluation factors that determine the success of a website [2].

2 Usability

In literature, usability is defined in several ways. According to ISO 9241 [3], usability defines “the extent to which a system, product or service can be used by specified users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specific usage context. In the literature, usability is defined as the “ability to be used” referring to a device [4] and depends on what the user wants to do [5]. According to Nielsen [6], usability “is a quality attribute that assesses how easy user interfaces are to use”, comprising 5 components: Learning: how easy it is for users to perform basic tasks the first time they encounter the design? Efficiency: Once users have learned the design, how can they perform tasks quickly? Memorability: When users return to design after a period of non-use, how easily can they re-establish themselves? Errors: How many errors do users make, how serious are those errors, and how easily can they recover from errors? Satisfaction: How enjoyable is it to use the design?

Usability is an important factor in developing successful interactive software applications. Furthermore, it is the most used concept in the area of software engineering [7]. Usability is a component used to define the quality of the app’s display and the quality of the interaction between users and the app. The benefits of usability are (i) reduce training time and cost, (ii) reduce errors that happened during interaction with the system, (iii) increase learning speed and satisfaction, (iv) improve quality and the interaction of users [8].

The studies described in the literature [9] indicate that usability is the most important parameter that affects the quality of websites evaluated by its users. According to Paplauskaite [10], the website’s usability determines its readability, intuitiveness, and comfort of use.

3 Methodology

This research aimed to evaluate, from the point of view of usability, the websites of hotel establishments in Terras de Trás-os-Montes. The Terras de Trás-os-Montes is a group of municipalities, made up of nine municipalities of the NUT III of Alto Trás-os-Montes, namely Alfândega da Fé, Bragança, Macedo de Cavaleiros, Miranda do Douro, Mirandela, Mogadouro, Vila Flor, Vimioso and Vinhais.

To this end, a search was carried out in the National Tourism Register (<https://registos.turismodeportugal.pt/>), by Tourist Enterprises, NUTS III—Terras de Trás-os-Montes, for the typology “Hotel Establishments” (Hotels, Apartment Hotels and Inns).

Table 1 Weights and measurements used to assess the usability of websites

Indicator	Weight (%)	Measure
Consistent design and navigation	10	Semaphore
Adaptive Design (Responsive)	10	Semaphore
Charging up to 2 s	10	Semaphore
Unique and immutable URLs	10	Semaphore
“Back” button functionality	10	Y/N
No invalid/broken links	10	Y/N
Links that change color when visited	10	Y/N
Readable font size	10	Y/N
Access with or without www (or https)	10	Y/N
Overall Performance Score	10	Semaphore

Data collection was performed by direct inspection of hotel websites. The evaluation was carried out by a group of three evaluators. Each analysis lasted an average of 10 min and was performed according to the list of indicators presented in Table 1. The list of indicators for evaluating the usability of websites was adapted from the study [11].

Each indicator has the same weight in the website’s evaluation, and the measures used were of the Yes/No and Semaphore types. For measures of the Yes/No type, a percentage of one hundred (100%) is attributed when an indicator is verified and zero (0%) when it is not verified. In the semaphore type indicators, the green check contributes the total weight of the indicator (100%), the yellow half (50%), and the red zero (0%).

4 Data Analysis and Results

4.1 Analysis and Indicators

Direct inspection of the websites of hotel establishments was carried out during the month of January 2021. Each of the indicators in Table 1 was analyzed by three researchers through direct observation and for some indicators, automatic tools were used, identified in Table 2.

The Screenfly tool was used for the analysis of the Adaptive Design (responsive) indicator and allows viewing on various devices, including 24” monitors and televisions. It can also customize screen measurements to any pixel value.

The PageSpeed Insights tool was used to analyze the Load up to 2 s, Readable font size, and Overall Performance Score indicators. This tool performs a series of tests on websites providing various indicators. For the analysis of the indicator No

Table 2 Tools used for automatically analyzed indicators

Indicator	Analyze	Tool
Adaptive design (Responsive)	Automatic	Screenfly
Charging up to 2 s	Automatic	PageSpeed insights
No invalid/broken links	Automatic	W3C link checker
Readable font size	Automatic	PageSpeed insights
Overall performance score	Automatic	PageSpeed insights

invalid/broken links, the W3C Link Checker tool was used, which automatically checks if all the links present on the website work correctly.

4.2 Data Sample and Its Characterization

Data on hotel establishments were collected from the National Tourism Register. Of the 31 hotel establishments, only 22 have websites, which were analyzed.

Of the 22 hotel establishments, 21 are Hotels, being only one of the inn typology. Of the hotels analyzed, three have 1 star, eight have 2 stars, seven have 3 stars and three have 4 stars, as shown in Table 3. They are mostly small, with the exception of Hotel Turismo São Lázaro, which has a capacity for 512 people (Table 3).

4.3 Website Usability Analysis

Of the 22 hotels with a website, six of them (Jorge V, Vinhais Hotel, Residencial Globo Lda., Hotel S. Roque, Hotel Estalagem Turismo, and Hotel Mirafresno) are not accessible, which demonstrates their lack of maintenance. The Ibis Bragança hotel does not have a local website, but a global website for the Ibis hotel chain, as such, it was excluded from the study.

The Usability criterion presents 10 indicators on the user experience to browse the hotel website. Aspects such as loading time of the website, consistent and adaptable design, broken links, among others, were evaluated.

The first indicator to be tested was Consistent design and navigation. Of the 15 websites evaluated, only Hotel Classis and Hotel Parador Santa Catarina did not score on this indicator (red semaphore), as it is a single page, with no possibility of browsing. All other websites got the highest score (green semaphore).

The adaptive design indicator (responsive design) has been checked to assess whether the websites of the evaluated hotels can be consulted through different devices, with different screen dimensions. Of the 15 websites analyzed, 10 comply with this requirement and have versions adaptable to different screen sizes.

As for the loading speed of all websites pages evaluated were fully loaded in less than 2 s, as such, they all obtained the maximum rating in this indicator.

Table 3 Characteristics of the analyzed hotel establishments

Name	Typology	Category	Capacity	Location
Jorge V	Hotel	*	46	Mirandela
Baixa Hotel	Hotel	**	32	Bragança
O Mirandês	Hotel	***	40	Miranda do Douro
VINHAIS HOTEL	Hotel	**	44	Vinhais
A Vileira	Hotel	**	24	Vimioso
Hotel Cabeço do Forte Miranda do Douro	Hotel	**	45	Miranda do Douro
Hotel Alendouro	Hotel	**	46	Macedo de Cavaleiros
Pousada de São Bartolomeu	Inn		56	Bragança
Hotel Turismo Miranda	Hotel	***	58	Miranda do Douro
Hotel Tulipa	Hotel	**	56	Bragança
Residencial Globo, Lda	Hotel	*	55	Mirandela
Hotel Póvoa de Além Sabor	Hotel	***	32	Vila Flor
Hotel S. Roque	Hotel	**	64	Bragança
Hotel & Spa Alfândega da Fé	Hotel	****	50	Sambade
Hotel O Encontro	Hotel	***	55	Sendim
HOTEL ESTALAGEM TURISMO	Hotel	****	52	Bragança
Hotel Classis	Hotel	**	40	Bragança
Hotel Parador Santa Catarina	Hotel	****	24	Miranda do Douro
Hotel Turismo São Lázaro	Hotel	***	512	Bragança
Hotel do Planalto	Hotel	*	64	Miranda do Douro
Hotel Mirafresno	Hotel	***	52	Miranda do Douro
Hotel Ibis Bragança	Hotel	***	126	Bragança

* - One Star

** - Two Stars

*** - Three Stars

With regard to unique and immutable URLs, 14 websites obtained a maximum rating and one a null rating, since the addresses used on the pages allow the identification of the hotel establishment to which they refer.

Regarding the functionality of the “Back” button, none of the analyzed websites has any problems.

Regarding the absence of invalid/broken links, 8 websites have at least one inconsistency in this requirement, which is why they have a null rating.

The indicator “links that change color when visited” was the worst indicator of all. None of the analyzed websites have this functionality, all having obtained a null score.

The readable font size was also analyzed, an indicator in which all analyzed websites had the maximum score.

Table 4 Average value for each of the indicators

Indicator	Average value (%)	Measure
Consistent design and navigation	86.67	Semaphore
Adaptive Design (Responsive)	66.67	Semaphore
Charging up to 2 s	100.00	Semaphore
Unique and immutable URLs	93.33	Semaphore
“Back” button functionality	100.00	Y/N
No invalid/broken links	46.67	Y/N
Links that change color when visited	0.00	Y/N
Readable font size	100.00	Y/N
Access with or without www (or https)	100.00	Y/N
Overall performance score	53.33	Semaphore

With regard to the indicator “Access with or without www (or https)” it was found that all analyzed websites allow access via the address in both forms (with or without the www or https).

Finally, for the “Overall Performance Score” indicator, the average performance between desktop and mobile was calculated. According to PageSpeed Insights’s assessment, between 0 and 49 is red (corresponding to zero), between 50 and 89 is yellow (corresponding to 50%) and between 90 and 100 is green (corresponding to 100%). On all websites evaluated, performance for desktop is always better than for mobile.

Of the analyzed websites, only 3 have a maximum overall performance, 2 have the minimum value and the remaining 10 have the intermediate value.

In Table 4 shows the average values obtained for each of the indicators.

There are four indicators with the highest score (Charging up to 2 s, “Back” button functionality, Readable font size, and Access with or without www (or https)) and one indicator with the lowest score (Links that change color when visited), meaning none of the sites evaluated comply with this indicator. Indicator No invalid/broken links also fails to obtain a positive score. The remaining indicators all score positively.

5 Conclusions

In this paper, we analyzed the websites of hotel establishments in Terras de Trás-os-Montes, according to the usability metrics.

Usability is an important parameter that clearly affects the quality of a website. It is important to make an investment so that all indicators are rated positively.

Of the analyzed Hotels's websites in the Terras de Trás-os-Montes, the global evaluation varies between 65 and 85%, although the evaluation is not excellent, it is not bad either. It is important to reflect on the less positive indicators so that the overall performance is improved.

At a time that mobile is increasingly used, it is important to understand how it affects the overall performance, since the analysis carried out showed that the Overall Performance Score was impaired due to poor performance at the mobile level.

Acknowledgements UNIAG, R&D unit funded by the FCT—Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education. UIDB/04752/2020

References

1. Teixeira, A., Fonseca, M., Castro, L., Garcia, J., Igreja, C., Costa Santos, C.: Escala Webqual na avaliação de websitesde hotéis do Porto. *Eur. J. Appl. Bus. Manage.* 55–72 (2018)
2. Aziz, N.S., Kamaludin, A., Slaiman, N.: Assessing web site usability measurement. *Int. J. Res. Technol.* 2(9), 386–392 (2013)
3. I.9241-220:2019, Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems. [Online]. Available: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en>. Accessed 30 Jan 2021
4. Quiñones, D., Rusu, C.: How to develop usability heuristics: a systematic literature review. *Comput. Stand. Interfaces* 53, 89–122 (2017)
5. Inostroza, R., Rusu, C., Roncagliolo, S., Rusu, V.: Developing SMASH: a set of smartphone's usability heuristics. *Comput. Stand. Interfaces* 46, 40–52 (2016)
6. Nielsen, J.: Usability 101: Introduction to Usability. 03 Jan 2012. [Online]. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>. Accessed on 01 Feb 2021
7. Grīnberga, S.: Software usability: concepts, attributes and associated health problems. *Proc. Latvian Acad. Sci.* 70(5), 266–268 (2016)
8. Seraj, M., Wong, C.Y.: A study of user interface design principles and requirements for developing a mobile learning prototype. In: *International Conference Computer and Information Science*, pp. 1014–1019 (2012)
9. Khalid, H., Hedge, A., Ahram, T.: *Advances in Ergonomics Modeling and Usability Evaluation*. CRC Press (2011)
10. Paplauskaitė, L.: Usability and usefulness in UX web design, [Online]. Available: <https://bitzesty.com/2014/05/15/usability-and-usefulness-in-ux-web-design/>. Accessed 28 Jan 2021
11. Sampaio, M., Alturas, B., Fernandes, A.L.: Avaliação de websites de Hotéis do ponto de vista do marketing digital. *Revista Ibérica Sistemas Tecnologias Informação* 2020(E34), 48–63 (2020)

Quality Factors for Agro-touristic Websites—An Exploratory Study



Karina Cicha  and Paulina Rutecka 

Abstract In this paper, the authors analyzed 6 selected websites of Polish agri-tourism farms using the WebQual questionnaire to determine their quality factors. The conducted analysis involved remote study with users' participation. The factors, which were taken under consideration during the analysis, focused on the functionality, usability, and aesthetics of the websites. The analysis was conducted according to WebQual method and used its questionnaire form. The main objective of the paper is to answer the question about the quality of the websites of agritourism farms in Poland.

Keywords Agritourism · Quality factors · Website quality · Website analysis

1 Introduction

Having a website at the beginning of the twenty-first century indicated that the company is innovative and strives to gain a competitive advantage [1]. Currently, although a basic website can be built even on its own with the use of free wizards or tutorials, many companies still do not use this option [2]. For enterprises and institutions, websites most often play a promotional role, they are used to present information or provide services [3–5]. They reflect the vision of the organization and its face in cyberspace [6], and can even be treated as the existence of an online institution [5]. They allow to reach a wider market and operate on a global scale [6]. They play an essential role in supporting business-to-consumer transactions [7], but only if they are trusted and user-serviceable.

According to numerous studies, it can be concluded that aesthetic, safe, and user-friendly websites with a simple interface attract more customers than websites that do not meet these requirements [3, 7, 8]. Having the ability to obtain information

K. Cicha (✉) · P. Rutecka

Department of Informatics, University of Economics, 40-287, Katowice, Poland

e-mail: karina.cicha@ue.katowice.pl

P. Rutecka

e-mail: paulina.rutecka@ue.katowice.pl

from many sources, the user must feel comfortable reading the presentation of the company's content, which should be well-organized and up-to-date [9]. A well-designed website can make customers remember the company's brand and have a more favorable attitude toward its products [10]. The quality of the website has a significant impact on the success of the organization and the business performance of the enterprise [11]. Research shows that when evaluating a site, customers also pay attention to factors such as trust and empathy [12, 13], navigation, security, economy, telepresence, and reputation [14].

There is little research on the quality assessment of agritourism farms' websites. Andreopoulou identified optimal websites and their features for agritourism facilities in Crete using the multi-criteria PROMETHEE method [15]. Król analyzed the visibility of agritourism websites in Visegrad Group countries, showing that agritourism enterprises in these countries often do not have websites or their quality is insufficient [16]. Beldona and Cai analyzed the effectiveness of agritourism websites in the USA, showing their low promotional value [17]. Dimitriadou presented a comparison of the adaptation to new technologies of exposure and promotion of websites of agritourism enterprises in Central Macedonia [18]. Other related studies concerned small tourist facilities, however, they were neither agritourism facilities nor located in rural areas.

The theoretical goal is to determine the quality factors of the websites of the agritourism farms. The practical aim of this study is to analyze the quality of the websites of agritourism farms in Poland, using the WebQual assessment method. The authors want to find out whether the websites of agritourism farms in Poland meet the quality factors. In addition, the results of the study can contribute to understanding how to increase the competitiveness of agritourism through website optimization.

The following research questions are proposed for this article:

RQ1: Do agritourism farms' websites meet the quality factors for websites?

RQ2: What are the most important quality factors of the agritourism farms' websites for the users?

The paper is organized as follows. Section 2 contains a review of the relevant literature on websites quality factors. Section 3 includes the methodology for data retrieval and processing, while Sect. 4 presents the data and quantitative results. In Sect. 5, the authors highlight the contribution of the research, discuss its limitations and, finally, draw conclusions about the results and propose possible future research avenues.

2 Literature Review

The hotel industry is extensive. Customers can choose from tens of thousands of accommodation facilities, which differ in terms of location, offer, and standard. This

means that finding a suitable place to rest is a difficult task for the client. According to numerous studies, tourists most often obtain information about a hotel facility via the Internet [19–21], because it reduces time and provides information in an easy way [22]. To make choosing easier, travel objects comparison websites and online travel agencies began to appear, operating in the commission model. Presence on these portals allows hotels owners to reach a wide group of customers. However, the commissions are often high [23] and information about the places is not extensive and complete.

Moreover, according to the Ipsos MediaCT 2014 study [19], tourists prefer to use search engines such as Google or Yahoo, rather than travel search engines [20]. For entrepreneurs, websites allow to reach a wider market and do business on a global scale [21], by playing a fundamental role in supporting business-to-consumer transactions [22]. The websites for enterprises are, in fact, a business platform, not just the channel for marketing [24]. Currently, just having a website is no longer sufficient for effective communication with the client. It must be a website that evokes a positive association with the company's brand. A high quality website can help reduce costs [24, 25] by becoming the main sales channel [21, 26], and its high quality and up-to-date content [25] can contribute to getting high SERP rankings and make it easier to find in search engines such as Google or Yahoo.

Having an individual website also increases the possibility to present the opinions of other tourists about the place, which is an important element in the process of making a purchase decision [27]. However, many websites are ineffective and fail to achieve the intended business goals [12], while a well-designed website could lead to better recognition of the company's brand by customers and a more favorable attitude toward the brand's products or services [10]. Not professionally designed and not updated websites may harm the company's image [23] and not fulfill its purpose.

Poland is among the EU countries with the largest number of farms and these are mainly farms with a small area. Agricultural production on small farms is usually insufficient to support the farmer's family, which is why many of them decide to diversify their sources of income [24, 28, 29] by conducting additional tourist activities, such as agritourism which is a specific type of tourist activity. It is closely related to rural areas and agricultural activity [30]. This concept is narrower in meaning than rural tourism [30] because apart from the fact that agritourism facilities are located in the countryside, they must be run on a working farm [31]. Agritourism activity is one of the business models supporting the sustainable development of rural areas [32, 33] and contributing to their revitalization [28, 30, 34]. Agritourism farms are usually small enterprises, often family and seasonal ones [2]. Their nature results from the type of agricultural activity [28, 35], offering guests products of this activity [20]. Due to the small scale of business and seasonality, as well as high risk related to the profitability of agricultural activities, owners of agritourism farms are forced to look for savings, including for marketing purposes.

Attracting tourists to the countryside is not an easy task [32]. However, the growing importance of sustainable development and new global threats, such as epidemics, make agritourism more and more popular every year in Europe, China, and the USA. The COVID-19 pandemic seemed to be an ideal moment for the development of

agritourism farms, which, as intimate places, away from large human clusters, could provide a sense of security. To attract new customers who were not previously interested in this form of recreation, owners of agritourism farms should place particular emphasis on promoting their tourist facilities on the Internet. Websites are the relatively cheap promotional solution [15] that allows a company to offer goods and services on a large-scale and, apart from selling services in the form of accommodation and meals, to show the natural and cultural elements of its community [32], thus attracting tourists' way to the countryside. This is especially important in the case of entities such as agritourism farms, where the website is most often the only way to familiarize the customer with the services offered before purchasing them because the quality of the accommodation cannot be checked before the consumption of its services [21, 36, 37]. Therefore, it is important to develop effective methods that can help companies evaluate the performance of their websites [24].

Scientists are constantly looking for website assessment methods that would be best suited to the goals and type of website, and at the same time convenient to use from the users' point of view [38]. The quality assessment answers the question of how well the system, in this case, the website, meets the needs of customers [11]. There is no single, universal measure of website quality, which results from the multitude of criteria and the variety of existing websites. This diversity results not only from the technology used but also from the goals to be achieved by the website. It was also observed that customers have different preferences for the same functionalities and page factors depending on the type of site [14] and the industry in which the company operates. Therefore, in the analysis conducted by the authors, the term website quality factors are used.

3 Research Method

Preliminary identification of the research group of agritourism farms was carried out in June 2020. The initial group consisted of 1925 tourist facilities from Poland, the addresses of which were in generally accessible databases. These facilities were asked via email whether they would agree to participate in the study. One hundred and seven objects agreed, which was the first stage of a large-scale study prepared as part of a doctoral dissertation. Of these, 76 have had active websites, which was the second selection criterion.

Then the basic technical parameters of the websites were checked: usage of a content management system (53 websites), website responsiveness (61 websites), securing the domain with an SSL certificate (33 websites), and usage of the Google analytics system (34 websites). The website's eligibility and approval for further research were verified and obtained in October 2020. Seven agritourism farms, out of those that used the CMS and monitored website traffic, finally agreed to participate in the quality survey. The websites of these farms were then tested by users.

The WebQual (eQual) ase don 4.0 [13] method, was used to assess the quality of agritourism farm websites in users' research. It is a survey method. The standard WebQual 4.0 tool contains 23 questions [39] divided into three dimensions of website quality: usability, information quality, and service interaction quality [13, 20]. The WebQual model is an ase don of the Servqual method, previously widely used to measure the quality of services, from which the criteria for user interaction with the website were taken [40]. Users evaluate such elements of the analyzed website as design, ease of use, navigation, attractiveness [41] and the quality of information, but it does not take into account such elements as transaction processes or the ase do's image [12]. WebQual is one of the best-formalized models, which is characterized by high versatility. It has been successfully used to evaluate e-commerce, e-government, and university websites [40]. The results are analyzed using the WebQual index, which takes into account the user's assessment of the criterion itself as well as its importance [13].

The group of respondents were people between the ages of 20–22, living in Poland, both men and women, students in the field related to tourism, without previous experience in such research, with no broad IT knowledge (basic Internet users). The selection of the study group was purposeful. At the beginning, the respondents were asked to determine the weight for each website feature on a 1–7 point scale. The questionnaire conducted in Polish contained 23 questions ase don Andry, Christianto, and Wilujeng's research [39].

4 Data and Results

4.1 Data Description

In the first step of the WebQual analysis, the importance of each criterion was established. The importance of the criterion was calculated as the arithmetic mean of the weights given by 6 respondents. Table 1 shows the ranking of means and standard deviations with an unbiased estimator for the questionnaire data [$n = 6$] of criteria importance. The most important quality criteria from the users' point of view are information security, information accuracy, information believability, and information timelessness. The least important quality criteria are the sense of community and a good reputation of a website.

Table 2 presents the responses of the respondents (the mean and standard deviation without the estimator bias for each site). The sites are marked successively as A1–A7.¹ The websites achieved the best average results for information believability (average 5.76) and learnability (average 5.60). The worst were marked sense of community (average 3.90) and user's experience provided with the website (average 4.81). The biggest discrepancies in the assessment of sites (the highest standard

¹ The WebQual survey in Polish along with the URLs of surveyed websites is available on: <https://forms.gle/vSTrsPfy1x4UPknc9>.

Table 1 Means and standard deviations for the questionnaire data [$n = 6$] of criteria importance

User satisfaction criteria [WebQual question]	Mean	Standard deviation
Easy to learn to operate [UQ01]	5.83	0.75
Interaction with the site is clear and understandable [UQ02]	6.33	0.82
The site is easy to navigate [UQ03]	6.50	0.84
The site is easy to use [UQ04]	6.50	0.84
Attractive appearance [UQ05]	5.33	1.03
Appropriate to the type of site [UQ06]	5.33	0.82
Conveys a sense of competency [UQ07]	6.00	0.89
Positive experience [UQ08]	5.33	1.21
Provides accurate information [UQ09]	6.67	0.82
Provides believable information [IQ10]	6.67	0.82
Provides timely information [IQ11]	6.67	0.82
Provides relevant information [IQ12]	6.50	0.84
Easy to understand information [IQ13]	6.00	0.89
Information at the right level of detail [IQ14]	5.67	1.03
Information in the appropriate format [IQ15]	5.33	1.03
Good reputation [SI16]	5.17	0.98
Safe to complete transactions [SI17]	5.83	1.17
The information feels secure [SI18]	6.83	0.41
Sense of personalization [SI19]	5.83	0.98
Sense of community [SI20]	3.50	1.38
Communicate with the organization [SI21]	5.50	1.05
Confident that goods/services will be delivered as promised [SI22]	5.50	1.76
Overall view of the Web site [SI23]	5.33	1.03

deviation) caused the question concerning the sense of community, while the highest agreement of answers concerned the area of ease to navigate.

Users rated (average higher or equal to 6.0) as providing believable information (IQ10) 4 websites (A1, A2, A6, A7). In users' opinion, 3 of the surveyed websites (A2, A6, A7) are easy to navigate (UQ03), and 3 (A1, A2, A6) are easy to use (UQ04). The websites A1, A2, and A7 are, according to users easy to learn to operate (UQ01), convey a sense of competency (UQ07), are appropriate to the type of site (UQ05), and bring the feeling of information security (SI18). None of the surveyed websites did not score higher or equal average than 6.0 in fields such as sense of personalization (SI19), sense of community (SI20), overall view of the website (SI23), and positive experience (UQ08). Only one website (A6) was marked as providing information at the right level of detail (IQ14), one (A2) as providing relevant information (IQ12), and one (A7) as appearing attractive (UQ05).

Table 2 Means and standard deviations (Dev.) for the questionnaire data [$n = 6$]

Web	A1		A2		A3		A4		A5		A6		A7	
	Mean	Dev	Mean	Dev	Mean	Dev	Mean	Dev	Mean	Dev	Mean	Dev	Mean	Dev
UQ01	6.00	0.89	6.50	0.55	4.83	1.17	5.17	0.75	4.83	1.33	5.83	0.75	6.00	0.89
UQ02	6.00	0.63	6.50	0.84	4.83	1.47	5.17	0.41	4.33	1.21	5.50	1.05	5.83	0.98
UQ03	5.67	1.21	6.33	1.03	4.67	1.21	4.67	0.82	4.33	1.37	6.00	0.89	6.00	1.10
UQ04	6.00	0.89	6.33	0.82	4.50	0.84	4.33	0.82	4.67	1.51	6.17	0.98	5.67	1.21
UQ05	5.67	1.97	5.83	1.17	3.17	1.33	4.17	1.17	3.67	1.75	5.67	1.37	6.00	1.10
UQ06	6.00	1.26	6.17	0.75	3.33	1.75	4.17	0.75	4.00	2.19	5.67	1.21	6.00	0.89
UQ07	6.00	1.26	6.00	1.55	2.83	2.14	4.33	1.03	4.33	1.51	5.83	1.17	6.17	0.98
UQ08	5.33	1.21	5.67	0.82	3.67	1.03	3.67	0.82	4.17	1.83	5.50	1.05	5.67	1.51
UQ09	5.67	1.21	6.33	0.52	4.50	1.38	4.67	0.82	4.17	1.72	6.33	0.82	5.67	0.82
IQ10	6.17	0.75	6.50	0.55	4.83	1.72	5.17	0.75	5.17	0.75	6.00	0.89	6.50	0.55
IQ11	4.67	1.51	6.17	0.75	5.00	1.26	4.83	0.75	5.67	1.03	6.00	1.26	6.00	0.63
IQ12	5.50	1.22	6.00	0.63	4.50	1.76	4.83	0.75	4.67	1.37	5.50	1.22	5.50	1.05
IQ13	5.67	1.03	6.17	0.41	5.17	1.72	5.33	0.52	4.33	1.37	5.67	1.03	6.17	0.75
IQ14	5.17	1.33	5.83	0.75	4.50	1.38	5.17	0.41	5.00	1.79	6.00	0.63	5.83	1.17
IQ15	5.67	1.03	6.50	0.55	4.33	0.82	5.17	0.75	4.33	2.25	5.33	0.82	6.17	0.98
SI16	5.17	1.17	5.50	0.55	4.33	1.03	4.83	0.75	4.33	0.82	5.67	0.82	6.00	1.26
SI17	6.17	1.17	6.17	1.17	5.00	0.89	5.00	1.10	5.00	1.67	5.50	1.05	5.67	1.21
SI18	6.33	0.82	6.17	0.75	4.67	1.03	5.00	0.63	4.67	1.21	5.67	0.82	6.00	0.89
SI19	5.50	1.05	5.67	0.52	4.67	1.51	4.50	0.55	4.33	1.86	5.67	1.03	5.83	0.98
SI20	3.50	1.52	4.17	1.72	3.50	1.87	3.83	1.47	3.83	1.72	4.33	1.86	4.17	1.72
SI21	5.67	0.82	6.50	0.55	4.17	1.33	5.33	1.03	5.00	1.41	5.67	1.03	5.83	0.75
SI22	5.50	1.38	6.67	0.52	4.83	1.17	4.67	1.37	5.00	1.26	5.67	0.52	6.17	0.75
SI23	5.33	1.21	5.67	0.52	3.83	0.98	4.17	0.98	4.50	1.87	5.33	0.82	5.83	1.17

Table 3 shows the weighted results, which, according to the WebQual method, show the differences in the priorities of users. Each respondent's rating of a website is multiplied by the weight they assign to the question. The mean of these results was then calculated. The highest score that a website can achieve is the value of the average weight of a given question multiplied by 7. In addition, to better present the results, Table 3 also shows the value of the WebQual Index [WQI] for the individual questions, i.e., the percentage that a website obtained in relation to the maximum value. The A2 website achieved the highest results, the A7 website placed second. Both of these websites achieved high WQI scores in the following areas: learnability, design adequacy, professional performance, information reliability, data security.

When assigning weights to questions, users identified which aspects they felt were most important. The weighted results show that in users' opinion, the aspect of information security (SI18) with an average weight of 6.83 is met by 3 pages: A1, A2, A7, achieving a WQI result of more than 0.85. Podczas nadawania wag dla pytań, użytkownicy określili, które aspekty są ich zdaniem najważniejsze. Z wyników ważonych wynika, że w opinii użytkowników aspekt Information feels secure (SI18) o średniej wadze 6.83 spełniają 3 strony: A1, A2, A7 osiągając wynik WQI POW. 0.85. Providing believable information (IQ10) was in second place, which in users' opinion were met by 4 websites (A1, A2, A6, A7), achieving the WQI score of more than 0.85. Equally important to the users (average weight 6.67) were providing timely information (IQ11), met by 3 websites (A2, A6, A7), with WQI score above 0.85 and providing accurate information (UQ09), met by 2 websites (A2 and A6). According to users, the ease to navigate (UQ03) and the ease to use (UQ04) were met by 3 websites, while providing relevant information (IQ12) only one.

4.2 Results

The analysis of the respondents' answers shows that for the agritourism farms websites users the most important features of the websites are those strictly related to the information on the website. First of all, the users need to feel that the information is secured. Also, the features of information itself, such as accuracy, believability, and timelessness are important to the users. In terms of websites functionality, the most important features for the users are ease of navigation, as well as ease of usage.

The analysis also revealed which features of the websites are not important for the users. The less important feature is providing the users with a sense of community. Users are also not interested in the good reputation of the website itself. Also, the information format and its adequacy to the type of the website seem irrelevant to the users. Surprisingly, the users also did not indicate a positive experience and overall view of the website as relevant features.

Table 3 Weighted scores (Wgt) and the WebQual index [$n = 6$]

Web Quest	Max. score	A1		A2		A3		A4		A5		A6		A7	
		Wgt	WQI	Wgt	WQI	Wgt	WQI	Wgt	WQI	Wgt	WQI	Wgt	WQI	Wgt	WQI
UQ01	40.83	35.33	0.87	38.00	0.93	28.67	0.70	30.00	0.73	27.83	0.68	33.67	0.82	35.33	0.87
UQ02	44.33	38.17	0.86	41.17	0.93	29.83	0.67	32.50	0.73	27.00	0.61	35.17	0.79	37.17	0.84
UQ03	45.50	37.50	0.82	41.17	0.90	30.00	0.66	30.17	0.66	27.83	0.61	39.17	0.86	39.67	0.87
UQ04	45.50	39.33	0.86	41.17	0.90	29.00	0.64	28.50	0.63	29.67	0.65	39.83	0.88	36.83	0.81
UQ05	37.33	31.00	0.83	30.67	0.82	16.00	0.43	21.67	0.58	19.33	0.52	30.67	0.82	32.00	0.86
UQ06	37.33	32.33	0.87	33.17	0.89	17.33	0.46	22.50	0.60	22.33	0.60	30.67	0.82	32.00	0.86
UQ07	42.00	36.50	0.87	36.17	0.86	16.17	0.38	25.67	0.61	25.67	0.61	34.67	0.83	37.17	0.88
UQ08	37.33	28.17	0.75	31.00	0.83	19.33	0.52	19.50	0.52	22.17	0.59	29.67	0.79	29.50	0.79
UQ09	46.67	38.33	0.82	42.33	0.91	29.83	0.64	31.00	0.66	27.50	0.59	42.33	0.91	38.00	0.81
IQ10	46.67	40.83	0.88	43.50	0.93	32.17	0.69	34.17	0.73	34.50	0.74	40.33	0.86	43.50	0.93
IQ11	46.67	32.00	0.69	41.50	0.89	33.00	0.71	32.50	0.70	37.67	0.81	40.33	0.86	40.00	0.86
IQ12	45.50	36.17	0.79	39.33	0.86	28.67	0.63	31.50	0.69	30.17	0.66	35.83	0.79	36.17	0.79
IQ13	42.00	34.00	0.81	36.83	0.88	30.50	0.73	31.83	0.76	25.67	0.61	34.33	0.82	37.00	0.88
IQ14	39.67	29.50	0.74	32.83	0.83	25.17	0.63	29.17	0.74	28.67	0.72	34.00	0.86	33.50	0.84
IQ15	37.33	30.67	0.82	34.33	0.92	23.17	0.62	27.50	0.74	21.33	0.57	28.33	0.76	33.33	0.89
SI16	36.17	27.33	0.76	28.67	0.79	22.67	0.63	24.83	0.69	22.83	0.63	29.50	0.82	30.50	0.84
SI17	40.83	35.83	0.88	35.33	0.87	29.17	0.71	29.00	0.71	28.00	0.69	32.00	0.78	32.83	0.80
SI18	47.83	43.50	0.91	42.33	0.89	31.83	0.67	34.17	0.71	31.67	0.66	38.83	0.81	41.00	0.86
SI19	40.83	32.33	0.79	33.33	0.82	26.33	0.64	26.33	0.64	25.67	0.63	33.33	0.82	34.00	0.83
SI20	24.50	13.50	0.55	16.00	0.65	14.17	0.58	15.00	0.61	15.33	0.63	17.00	0.69	16.33	0.67
SI21	38.50	30.67	0.80	35.83	0.93	22.17	0.58	28.83	0.75	26.50	0.69	30.83	0.80	32.33	0.84
SI22	38.50	32.17	0.84	37.00	0.96	26.83	0.70	26.17	0.68	27.00	0.70	31.17	0.81	34.83	0.90
SI23	37.33	29.17	0.78	30.17	0.81	20.17	0.54	21.83	0.58	22.50	0.60	28.33	0.76	31.67	0.85
Total	939.1	764.3		821.8		582.1		634.3		606.8		770.0		794.6	

5 Conclusion and Discussion

In this paper, we presented an analysis of the quality factors of agritourism farm websites using the WebQual method. The findings of our study indicate that some websites of the analyzed farms meet some of the users' expectations showing at the same time in which aspects they could be improved. The low-rated websites were made a few years ago and are not updated frequently. Users, in particular young respondents, who are used to visiting modern websites, require the company to follow trends in order to feel safe and take advantage of the company's offer. The feeling of security and credibility of information, compliance of the presented information and graphic materials with the actual state of affairs is the most important for them. They definitely do not feel the need to be part of a community in order to benefit from the services of an agritourism farm. The users expect the website to be intuitive to use. The top-rated sites (A1, A2, A6, and A7) are extensive and provide information not only about the company's services but also about the area and local attractions which can use during your vacation.

Research using the WebQual method allows understanding the aspects of the website users pay attention to and how they evaluate them. However, the WebQual method can be difficult to apply to a larger number of websites because of the number of questions. Furthermore, the answer must be preceded by a user analysis of the page itself, which is time-consuming. The analysis performed for 7 websites required each respondent to vote in 184 questions. The solution to this problem for a larger number of websites and with a greater number of respondents could be assigning a few random pages for analysis to each respondent and correlating these responses, which, however, is subject to the risk of error.

The limitation of our study is the fact that only 7 websites of agritourism farms were analyzed, which is a small segment of all agritourism enterprises. Another limitation was the study on a small sample of respondents ($n = 6$). Moreover, the study concerned only with agritourism farms located in Poland.

The authors, however, plan to conduct another study in the near future, to develop an expert research framework based on technical data, and then to correlate the results of the expert analysis with the users' expectations. The framework could be an alternative to user research. It is also planned to conduct a survey with a larger number of respondents in order to achieve more representative results regarding the expectations and satisfaction of users.

The practical implication of this research is an indication to the owners of agritourism farms which elements of websites are important for users and which are marginal. The results of this study can help owners implement changes and apply the right technologies in their tourism marketing.

References

1. Brady, R., Forrest, E., Mizerski, R.: *Marketing w internecie*. Polskie Wydawnictwo Ekonomiczne, Warszawa (2002)
2. Rutecka, P.: Właściciel agroturystycznego przedsiębiorstwa rodzinnego jako brand-hero marki. In: *Nowe trendy w rozwoju firm rodzinnych*, pp. 79–93. Wydawnictwo Politechniki Częstochowskiej, Częstochowa (2019)
3. Fabisiak, L., Wolski, W.: Metody analizy wielokryterialnej w ocenie użyteczności serwisów internetowych. *Stud. Mater. Pol. Stowarzyszenia Zarządzania Wiedzą*. **27**, 116–125 (2010)
4. Radziszewska, A.: Kryteria oceny jakości komercyjnych serwisów internetowych. *Mark. Rynek*. 49–56 (2014)
5. Utami, I.S., Winarno, Setiadi, H.: Analysis the effect of website quality on user satisfaction with the WebQual 4.0 method and importance-performance analysis (IPA) (case study: SPMB Sebelas Maret University's Website). *J. Phys. Conf. Ser.* **1842**, 012003 (2021). <https://doi.org/10.1088/1742-6596/1842/1/012003>
6. Napitupulu, D.: Analysis of factors affecting the website quality (study case: XYZ University). *Int. J. Adv. Sci. Eng. Inf. Technol.* **7**, 792 (2017). <https://doi.org/10.18517/ijaseit.7.3.1748>
7. Fam, S.F., Huang, J., Chuan, Z.L., Khalil, S.N., Prastyo, D.D., Nusa, F.N.M.: Fuzzy TOPSIS method as a decision supporting system to rank Malaysia online shopping website quality during COVID-19 MCO 2020. *Int. J. Emerg. Trends Eng. Res.* **8**, 6397–6405 (2020). <https://doi.org/10.30534/ijeter/2020/239892020>
8. Hsu, M.H., Chang, C.M., Chu, K.K., Lee, Y.J.: Determinants of repurchase intention in online group-buying: The perspectives of DeLone & McLean is success model and trust. *Comput. Human Behav.* **36**, 234–245 (2014). <https://doi.org/10.1016/J.CHB.2014.03.065>
9. Rozekhi, N.A., Hussin, S., Noor, A.M.: Attributable E-commerce toward Purchase Intention: Online Search of Food Product. *SIJ Trans. Adv. Sp. Res. Earth Explor.* **4**, 6–12 (2016). <https://doi.org/10.9756/SIJASREE/V4I1/0203410401>
10. Bai, B., Law, R., Wen, I.: The impact of website quality on customer satisfaction and purchase intentions: evidence from Chinese online visitors. *Int. J. Hosp. Manage.* **27**, 391–402 (2008). <https://doi.org/10.1016/j.ijhm.2007.10.008>
11. Karabasevic, D., Stanujkic, D., Maksimovic, M., Popovic, G., Momcilovic, O.: An approach to evaluating the quality of websites based on the weighted sum preferred levels of performances method. *Acta Polytech. Hungarica* **16**, 195–215 (2019). <https://doi.org/10.12700/APH.16.5.2019.5.11>
12. Kim, S., Stoel, L.: Apparel retailers: website quality dimensions and satisfaction. *J. Retail. Consum. Serv.* **11**, 109–117 (2004). [https://doi.org/10.1016/S0969-6989\(03\)00010-9](https://doi.org/10.1016/S0969-6989(03)00010-9)
13. Barnes, S., Vidgen, R.: An integrative approach to the assessment of e-commerce quality. *J. Electron. Commer. Res.* **3**, 114–127 (2002)
14. Lee, Y., Kozar, K.A.: Investigating the effect of website quality on e-business success: an analytic hierarchy process (AHP) approach. *Decis. Support Syst.* **42**, 1383–1401 (2006). <https://doi.org/10.1016/j.dss.2005.11.005>
15. Lemonakis, C., Zopounidis, C., Koliousska, C., Andreopoulou, Z.: Internet and agro-tourism sector for regional development in Crete: a multicriteria ranking. *Int. J. Inf. Decis. Sci.* **9**, 116 (2017). <https://doi.org/10.1504/IJIDS.2017.10005872>
16. Król, K.: Promoting of agrotourism on the internet. A lesson from the Visegrad Group countries. *J. Agribus. Rural Dev.* **16**, 805–813 (2017). <https://doi.org/10.17306/J.JARD.2017.00383>
17. Beldona, S., Cai, L.A.: An exploratory evaluation of rural tourism websites. *J. Conv. Event Tour.* **8**, 69–80 (2008). https://doi.org/10.1300/J452V08N01_04
18. Dimitriadou, E., Lamprouli, I., Iakovidou, O.: Adapting new technologies of exposure and promotion for agro-tourism enterprises with comparative evaluations. *Asian J. Agric. Rural Dev.* **9**, 274–283 (2019). <https://doi.org/10.18488/journal.1005/2019.9.2/1005.2.274.283>
19. Ipsos MediaCT: Google travel study. The 2014 traveler's road to decision. (2014)

20. Oplanić, M., Težak Damijanić, A., Saftić, D., Čehić, A.: The internet as a source of information about local agro-food product. *J. Cent. Eur. Agric.* **20**, 759–769 (2019). <https://doi.org/10.5513/JCEA01/20.2.2127>
21. Rabadán-Martín, I., Aguado-Correa, F., Padilla-Garrido, N.: Facing new challenges in rural tourism: signaling quality via website. *Inf. Technol. Tour.* **21**, 559–576 (2019). <https://doi.org/10.1007/s40558-019-00157-y>
22. Heung, V.C.S.: Internet usage by international travellers: reasons and barriers. *Int. J. Contemp. Hosp. Manage.* **15**, 370–378 (2003). <https://doi.org/10.1108/09596110310496015>
23. Gössling, S., Lane, B.: Rural tourism and the development of Internet-based accommodation booking platforms: a study in the advantages, dangers and implications of innovation. *J. Sustain. Tour.* **23**, 1386–1403 (2015). <https://doi.org/10.1080/09669582.2014.909448>
24. Bršćić, K., Dropulić Ružić, M., Ružić, D.: Is Agrotourism Popular on the Internet? *J. Cent. Eur. Agric.* **13**, 221–235 (2012). <https://doi.org/10.5513/JCEA01/13.2.1038>
25. Duarte, P., Pais, A.R.: Use and perception of the internet as a marketing tool to promote rural tourism. *Work. Pap. Gestão, Econ. e Mark. (Management, Economic and Marketing Working Paper)* (2010)
26. Li, X., Wang, Y., Yu, Y.: Present and future hotel website marketing activities: change propensity analysis. *Int. J. Hosp. Manage.* **47**, 131–139 (2015). <https://doi.org/10.1016/j.ijhm.2015.02.007>
27. De Pelsmacker, P., van Tilburg, S., Holthof, C.: Digital marketing strategies, online reviews and hotel performance. *Int. J. Hosp. Manage.* **72**, 47–55 (2018). <https://doi.org/10.1016/j.ijhm.2018.01.003>
28. Akpınar, N., Talay, İ., Ceylan, C., Gündüz, S.: Rural women and agrotourism in the context of sustainable rural development: a case study from Turkey. *Environ. Dev. Sustain.* **6**, 473–486 (2005). <https://doi.org/10.1007/S10668-005-5633-Y>
29. Santeramo, F.G., Barbieri, C.: On the demand for agritourism: a cursory review of methodologies and practice (2016). <https://doi.org/10.1080/21568316.2015.1137968>
30. Sznajder, M., Przezbórska, L.: *Agroturystyka. Polskie Wydawnictwo Ekonomiczne, Warszawa* (2006)
31. Wojciechowska, J.: Systematyka i wyróżniki pojęć dotyczących turystyki na obszarach wiejskich. *Stud. nad Tur. Tradycje, stan obecny i Perspek. Badaw.* **5**, 297–310 (2007)
32. Karagiannis, S., Stavroulakis, D.: Development of agrotourism through the application of internet technologies: problems and perspectives. In: *CEUR Workshop Proceedings*, pp. 299–308. CEUR-WS (2011)
33. Fanelli, R.M., Romagnoli, L.: Customer satisfaction with farmhouse facilities and its implications for the promotion of agritourism resources in Italian municipalities. *Sustainability* **12**, (2020). <https://doi.org/10.3390/su12051749>
34. Busby, G., Rendle, S.: The transition from tourism on farms to farm tourism. *Tour. Manage.* **21**, 635–642 (2000). [https://doi.org/10.1016/S0261-5177\(00\)00011-X](https://doi.org/10.1016/S0261-5177(00)00011-X)
35. Stoddard, J.E.: *Sustainable Tourism The Case of the Blue Ridge National Heritage Area* (2018). <https://doi.org/10.1177/1938965508320574>
36. Casaló, L.V., Flavián, C., Guinalfú, M., Ekinci, Y.: Avoiding the dark side of positive online consumer reviews: enhancing reviews' usefulness for high risk-averse travelers. *J. Bus. Res.* **68**, 1829–1835 (2015). <https://doi.org/10.1016/J.JBUSRES.2015.01.010>
37. Litvin, S.W., Goldsmith, R.E., Pan, B.: Electronic word-of-mouth in hospitality and tourism management. *Tour. Manage.* **29**, 458–468 (2008). <https://doi.org/10.1016/J.TOURMAN.2007.05.011>
38. Chmielarz, W., Zborowski, M.: The selection and comparison of the methods used to evaluate the quality of e-banking websites: the perspective of individual clients. In: *Procedia Computer Science*, pp. 1903–1922. Elsevier B.V. (2020)
39. Fernandes Andry, J., Christianto, K., Wilujeng, F.R.: Using Webqual 4.0 and importance performance analysis to evaluate e-commerce website. *J. Inf. Syst. Eng. Bus. Intell.* **5**, (2019). <https://doi.org/10.20473/jisebi.5.1.23-31>

40. Ziemia, P.: Zastosowanie korelacji i metody wielokryterialnego podejmowania decyzji przy doborze kryteriów oceny serwisów internetowych. *Ekonometria* 154–164 (2011)
41. Piwowarski, M., Ziemia, P.: Zastosowanie modelu Equal 4.0 do oceny złożonych serwisów internetowych. *Zesz. Nauk. Uniw. Szczecińskiego. Ekon. Probl. Usług.* 681–688 (2010)

Tracking Techniques in the Study of Tourists' Spatiotemporal Behavior



Irina Clara, Ricardo Correia , and Aida Carvalho 

Abstract The knowledge of tourism consumer behavior is essential for destinations and tourism organizations to improve their offer, acting as one of the main ingredients for the development of memorable experiences. The spatial–temporal tracking of tourists is one of the techniques associated with the knowledge of consumer behavior and which allows valuable information to be obtained. On the other hand, the demands of tourist consumers are increasing, which make it even more important to have a better knowledge of them. In this context, this study aims to present the traditional and modern methods of tracking tourists that could be used for the improvement of the tourism offer. To this end, a methodology of comprehensive literature review and state-of-the-art verification was adopted, to identify the main tracking techniques in the spatiotemporal behavior of tourists. It is concluded that these techniques should be used together, as they complement each other. Given the constant technological advances, this theme has the potential to be further developed in terms of literature, recommending future empirical research on the phenomenon.

Keywords Tracking technologies · Georeferencing · GPS · Tourism

1 Introduction

Methods such as direct observation and questionnaires have been used over the decades, to try to understand the spatial and temporal behavior of tourists. With the technological advances in this sector, new techniques have been developed

I. Clara · R. Correia (✉) · A. Carvalho
Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
e-mail: ricardocorreia@ipb.pt

I. Clara
e-mail: a48289@alunos.ipb.pt

A. Carvalho
e-mail: acarvalho@ipb.pt

R. Correia · A. Carvalho
CiTUR Guarda—Centro de Investigação, Desenvolvimento E Inovação Em Turismo, Faro, Portugal

that allow the tracking of tourists, which have been filling the gaps of traditional methods. Currently, various techniques are used to try to understand the behavior and motivations of tourists, with the aim of improving the services and experiences provided.

Tracking techniques help to understand consumer behavior patterns [1], that is, how they behave at the destination, to ensure that all attractions and services are properly managed, are not overcrowded and are capable to provide a positive experience for tourists [2]. A greater understanding of tourist movements, the factors influencing the relationships between destinations, and the spatiotemporal activities and behavior of tourists allow a greater understanding of the impact of tourism at different levels. Thus, it allows to develop an offer more oriented to specific segments and to implement management and marketing strategies more suited to a specific destination or attraction [2, 3].

Considering the relevance of tracking techniques for the management of tourist destinations and for the enhancement of experiences, as well as the technological evolution, this article intends to identify the main tracking techniques and highlight the main contributions of the most modern methods, through the collection of practical examples.

To achieve these goals, the article is organized as follows: In Sect. 2, the main tracking techniques are presented, listing their strengths and weaknesses. Subsequently, some examples of how georeferencing technologies can be used are presented with a summary table of some studies and respective results. The article ends with a set of conclusions and suggestions for future research.

2 Main tracking techniques: strengths and weaknesses

One of the oldest traditional methods is direct observation, which can be summarized as “identifying, following, observing, and mapping” [4, p. 143] and involves participatory observation and observation at a distance [5]. After direct observation, questionnaires and interviews began to be used, to fill in the gaps of observation methods. Later, Hartmann [6] used a remote observation technique, which consisted of placing a camera to understand the routes of tourists. This method had the major disadvantage of no longer being useful when tourists left the camera’s viewing angle [4]. Later, techniques such as travel diaries and photographic diaries were used [5]. The great weakness of traditional methods is that is impossible to understanding the motivations of tourists to visit certain attractions. Also, traditional methods are deficient with regard to the accuracy and validity of the data collected [4]. With the technological advances verified in the last decades, new tracking methods have emerged, which allow obtaining high-resolution information associated with tourist movements [5].

Modern methods and techniques include the use of global positioning system (GPS) devices, Bluetooth tracking, smart cards, mobile tracking system, smartphone applications (APP), and georeferenced information shared on social networks, among others [5, 7, 8].

Data obtained through GPS uses sequential tourist patterns on a local scale: number of locations visited, their order, and the time each tourist spent on a particular attraction [1]. The advantages of GPS technologies are the public availability of GPS devices (these can be used all over the world), as well as their availability in small markets, at a relatively low price. This is a mature technology, and the devices are easy to use and transport [5]. Furthermore, there is a great willingness of tourists to participate in questionnaires based on the use of GPS devices, since tourists do not feel that their privacy is being invaded [2]. GPS devices allow to collect large amounts of accurate information, in a short period of time and with a high resolution. However, this information may contain errors in certain cases, such as in closed areas or areas with a lot of infrastructure. Other disadvantages of these technologies are battery life, tracking accuracy in urban environments, and overall quality of tracked positions, due to signal loss in certain areas and the delay for the device to re-find signal, which can range from 5 to 30 s. Therefore, the quality of data collected is strictly dependent on the characteristics of the destination [2, 5]. Martins [5] also points out the high risk of losing the equipment, which is expensive.

Bluetooth analysis are based on systematic observations, which detect mobile devices connected to Bluetooth, making use of the digital footprint that visitors leave unconsciously, to classify the behavior of visitors through the journeys made and by the relationship of paths with the time spent on each work [1]. This method is suitable for smaller scales, such as a specific historic site, or closed places. The time spent in a particular location reflects the level of interest. This technique has advantages such as more detailed information, not requiring prior registration, nor using any additional devices, which leads to a larger sample. Also, Bluetooth proximity detection works inside buildings or near structures, where GPS connectivity is limited. The non-intrusive nature eliminates data bias that can exist if individuals know they are being tracked. These advantages make this method suitable for detecting the sequential movements of consumers between locations, without specifying their activities, attributes, or thoughts [1, 9]. The disadvantages are that the Bluetooth proximity sensor only provides the sequence of individual mobile device transitions between nodes, while the GPS tracks all the device's movements. On the other hand, the composition of the sample depends on the decision of device users to activate their Bluetooth [1].

Smart cards are easy to use, non-intrusive and have a low cost associated with a large sample size. However, this technique is limited to the attractions that the tourists visit [5]. Whenever a transaction is made, the smart card system records individual transactions for passenger entry and exit, creating trip logs. This data reveals detailed information about users' daily itineraries and information about the users themselves. This information is, for example, the identification number of the card, the station, the route, the method of transport used, the type of passenger, and the distance covered, for each individual user [10]. The main purpose of this method

is to collect revenue and increase passenger comfort, but this information generates a lot of accurate and detailed information for planning the transport route efficiently and to define behaviors and characteristics of tourists [13].

Mobile tracking systems are used to track the movements of tourists, through passive information that is generated each time a mobile phone communicates with a network, which can be associated with the information of the owner of the SIM card, allowing to associate the age and gender, for example, and thus make a spatiotemporal analysis of the movements of different groups [11]. The advantages of this method are high resolution, low cost, and that is not intrusive. On the other hand, it has disadvantages such as signal blocking in certain areas, problems related to privacy and security, and the impossibility to apply surveys [5].

The geographic information provided by social networks was also considered as methods to provide valuable information regarding the dynamics of human mobility, with a great spatiotemporal resolution. Social networks allow individuals to socially interact and share content about the places where they are, providing a source of information to understand and assess human behavior [11]. These user-generated data allows to comprehend mobility patterns, which contribute to a greater understanding of the activities that tourists perform during their trips, such as the temporal and spatial distribution of visitors, the points of interest of most tourists, the itinerary traveled, and the origin of the tourists [12]. Data provided by social networks can assist product development and destination management and can help to make more accurate predictions, as well as respond to demand effectively, as they provide a direct communication bridge between locations and tourists. These data represent an opportunity to improve knowledge about an environment, to detect flows of people and the use of services in a certain area. This knowledge allows to make better and more sustainable decisions. On the other hand, it allows a better understanding of the needs and desires of consumers, to adapt the tourist products of a particular destination. Social networks and the information they provide affect the image of destinations and may generate more tourist demand. The advantages are that data requires fewer financial resources and is more accurate, having fewer limitations in spatial and temporal analysis. On the contrary, one of the main criticisms of social media data is the biases that come from specific characteristics of platforms, such as their age, the geographical location, and the social class of its users. A solution to this problem could be the combination of different social networks as sources of information [11].

Although the different methods are effective and have advantages, the best results are produced when several methods are applied simultaneously, as they effectively complement each other [2].

3 Applications of Georeferencing Technologies

3.1 GPS

GPS devices are especially suitable for tourists traveling on cruises as there is only one entry/exit point and the visit time is relatively short. In this field, GPS technologies make it possible to collect and analyze precise information related to the experiences of passengers, obtaining information such as the attractions visited, the time spent in each attraction, the route taken, the type of transport used, among others. These results allow a better understanding of tourists' experiences, based on their levels of satisfaction, the relationships between specific attractions and the intention to revisit or recommend the destination. This information is important for better transport planning, for adjusting the information provided to tourists about accommodation or restaurants, for improving the management of tourist services and flows, and for providing a positive destination experience [2].

Other study used GPS to analyze the spatiotemporal behavior of tourists, considering the concentration of their movements and the attractions they visited, concluding which are the most visited attractions in the city of Lisbon. The study also investigated the movements of tourists, that is, the routes taken and the transport used, allowing to draw conclusions about the multi-attraction dimension of the city and the categories of most visited attractions [13].

3.2 Bluetooth

A study was carried out using Bluetooth technologies to analyze the behavior of tourists at the Louvre Museum [1]. One of the biggest problems for many museums is crowds and overcrowding (the number of visitors exceeds the space capacity). These factors reflect the attractiveness of the museum, resulting in a positive economic impact. However, these problems are difficult to control and the high number of visitors can also have negative effects, such as a reduction of the conditions and quality of visits, having a negative impact on the consumers' experience [1]. In the case of museums, georeferencing techniques, specifically through Bluetooth, allow to conclude the most visited pieces and the path taken by the visitors, that is, in which order they visit them. This knowledge can contribute to improving the layout and environment of museums (designing more appropriate spatial arrangements, helping to manage flows more efficiently and dynamically) and to improving the visitor experience [1].

3.3 *Smart Cards*

In the study made by Ali et al. [9], smart cards were used as an input to analyze the behavior of tourists in Seoul, South Korea, identifying segments based on information about the start and end of trips, duration, and means of transport used [10]. In another study, the authors used this method to identify different clusters of passengers in the Costa Dorada area, in Spain, according to their activities and their spatial profiles. In this way, transport authorities were able to detect weaknesses in transport networks and adopt specific measures to optimize the network. The identification of profiles allows for strategic, tactical, and operational management of public transport, by predicting the evolution of demand, identifying defects, or developing performance indicators [14].

3.4 *Social Networks*

Twitter is the most widely used platform in respect of tourism research, as it provides useful information about the tourist flows. On Twitter, users can choose to place their location in their publications, using GPS, which allows to carry out a spatiotemporal analysis. Even though only about 1% of publications are associated with a location, it is still a rich source of information, as there are several text extraction and analysis techniques that allow an interpretation of its content. However, it is important that the information is properly filtered [11]. In addition to tourist research, this social network is useful in tourism management, and a study even shows that Destination Management Organizations (DMOs) that actively use this social network to promote their products, achieve a higher rate of occupation, due to the number of shares and user responses [15]. On the other hand, Kovacv et al. [10] combined social networks with mobile data positioning to analyze international tourist flows in an urban center in Hungary. The authors analyzed Twitter publications and identified events that attracted foreign tourists, then used data from SIM cards to identify tourist segments. Based on the combination of social networks and mobile positioning data, they were able to identify the attractions that generated the most significant tourist arrivals in the city [11].

Barros et al. [11] used social network and GPS to understand the characteristics of tourists who visited the Teide National Park, in Santa Cruz de Tenerife, Spain, through information such as the points of interest, visitor concentration and movements, country of origin, and distribution of visitors over time [12]. Huertas and Martínez-Rolán [16] also carried out a study based on publications from the social network, this time from Instagram. Instagram stands out as a social network that allows to stimulate communication through photographs, videos, or experiences about a particular product or tourist destination [16]. The authors analyzed the publications of georeferenced photographs and videos, user reactions and created

an impact factor that measures the popularity of destinations among different users, showing the usefulness of an impact factor for DMO's [17].

3.5 Main Contributions of Modern Tracking Technologies

Several authors use modern methods of tracking tourists using GPS [2, 3, 12, 13, 18–20], Bluetooth [1, 9], mobile tracking system [11, 20, 21], smart cards [10, 14], smartphone applications [22–24], and georeferenced information shared on social networks [11, 12, 17, 25–29]. Some of these studies are summarized in Fig. 1.

Study	Methodology	Main Conclusions
Caldeira (2014)	GPS; Survey.	Evident contrasts in the territory of the study area, which confirm the selectivity of the tourist use of space: visits to attractions in the historic center of the city and in Belém, followed by Parque das Nações, the Sintra area and Cascais/Estoril; Most of the tourists made a complex route, which is in line with the preference of independent visitation, and frequently use public transport; Urban tourists generally direct their visit to multiple attractions, the most visited being heritage attractions and historic/urban areas.
Spangenberg (2014)	GPS; Survey.	Trend of movements in two directions: Schierke, Torfhaus, especially by first time visitors; Repeat tourists value walking, and use less frequented routes.
Versichele, de Groote, Claeys Bouuaert, Ncutens, Moerman & Van de Weghe (2014).	Bluetooth.	Heterogeneous visitor group, resulting from the combination of open and closed attractions; Only a small fraction of tourists combined more than one closed attraction; Visitors of more than one day: associations between the center and remote attractions; Hotel guests: differences from general visitors; Two locations with considerable associations: "City Museum of Ghent (STAM)" and "House of Alijn'.
Yoshimura, Sobolevsky, Ratti, Girardin, Carrascal, Blat & Sinatra (2014)	Bluetooth.	The behavior of visitors who stay less time and those who spend more time in the museum are similar in terms of works observed and routes taken (those who spend more time take longer to appreciate the same works); The number of nodules visited is independent of the duration of the visit; The frequency of nodules visited per hour is independent of the duration of the visit; The sequential movement of visitors is limited in terms of order and path taken, although there are several possibilities.
Viswanath, Yuen, Ku & Liu (2015)	Smartphone app; Survey.	The Smart Tourist application is able to locate all the places the tourist traveled through, as well as demarcate the locations according to the time spent at each location; Colac is an interesting village to visit.
Raun, Ahas & Tiru (2016)	Mobile tracking system.	Visitor patterns vary by geographic level; Visitors from Laane County are associated with those from Saaremaa, and those from Jogeva County are associated with those from Tartumaa; Different seasonal patterns; Differences in length of stay; Time preferences of different seasonalities (e.g. German visitors prefer summer, while Russians are insensitive to this season).

Fig. 1 Main contributions of modern tracking technologies

<p>Abarca (2016)</p>	<p>Georeferenced photos: Panoramio.</p>	<p>Distribution patterns of tourists' and residents' photographs showed both similarities and differences; Photographs of tourists showed greater spatial concentration than those of residents; Spatial clusters coincided with areas of the city that are known to attract more tourists; Tourists tend to concentrate in areas with the most tourist potential in the city; The most prominent areas were well demarcated, being frequently traveled together by tourists: Belém, the historical center and Parque das Nações. Tourists' photographs are a very relevant indicator of the perceived image of a destination.</p>
<p>Ferrante, De Cantis & Shoval (2018)</p>	<p>GPS; Survey.</p>	<p>Cruise passengers behave differently at destination and between destinations; Passengers with higher incomes, higher levels of education, and ages between 36 and 55, seek more intense experiences at destinations, in terms of time spent, attractions visited, and means of transportation.</p>
<p>Raun, Shoval & Tiru (2019)</p>	<p>GPS; Mobile tracking system.</p>	<p>Importance of the role played by accesses/entrances in the distribution of tourist flows within a country, since most tourist activity is located in its vicinity; In Estonia, the main access is Talin (dominance of the country's capital); In Israel, the main entrance is Ben Gurion airport; Most tourist activities are concentrated in urban areas near the entrances.</p>
<p>Huertas & Martínez-Rolán (2020)</p>	<p>Social networks: Instagram - photos, videos and stories.</p>	<p>Importance of taking into account the population density of the territories, when measuring their popularity through the publications obtained; Some destinations with lower population get more publications (Barcelona and Seville), which indicates that these destinations are more attractive for the publications, since they have a greater number of tourists more active on Instagram; Relevance of user interactions (likes, comments) to the publications for measuring popularity, as these are what really mirror the success of the publications; The Impact Factor on Instagram is a much more appropriate index to measure the popularity of destinations through the publications of photos and videos.</p>
<p>Gutiérrez, Doménech, Zaragoza & Miravet (2020).</p>	<p>Smart cards.</p>	<p>Clusters: sprawled, semi-sprawled, main cities, coastal; High spatial concentration of activity around the most tourist areas; Moderate volume of transactions in the main cities; Large difference in terms of proportion of the profiles; Sprawled cluster: significant presence at all bus stops; Semi-sprawled cluster: presence in main cities almost non-existent, more predominant at coastal stops, concentrated on few routes and visiting few attractions; Main cities cluster: more important in the main cities of the region; Coastal cluster: higher percentage of stops near the sea.</p>
<p>Kovács, Vida, Elekes & Kovalesik (2021)</p>	<p>Social networks: Twitter; Mobile tracking system.</p>	<p>Most mentioned events on Twitter primarily related to sports; Two specific events: "Rugby Europe Women's 7s Trophy" and "11th International Dragon Boat Federation Club Crew World Club Crew World Championships", which attracted players, rowers, managers, family members, fans and spectators; Significantly higher number of foreign tourists in the summer months, with another visible peak in December; Main foreign tourists nationalities: Germany, Canada, United States, Czech Republic, and Poland; Most visited places: downtown (concentration of attractions) and restaurant and accommodation areas; Movements in the city suburbs: more dispersed and more "hot spots".</p>
<p>Casado-Díaz, Navarro-Ruiz, Nicolau & Ivars-Baidal (2021)</p>	<p>GPS; Survey.</p>	<p>The behavior patterns of both types of cruise passengers (independent and guided) were different according to the numbers of nodes visited; The different mobility patterns associated with passengers' visits influenced their spending differently (independent passengers: higher spending); Passengers staying at a single node tended to spend more than those moving to multiple nodes.</p>
<p>Gunter & Onter (2021)</p>	<p>Social networks: Instagram.</p>	<p>The top 50 mostly comprise traditional attractions, followed by hotels, recreational areas, cafes/ restaurants, and events; The top 10 locations for residents and visitors are identical.</p>

Fig. 1 (continued)

4 Conclusions

A greater and better understanding of the behaviors and motivations of tourists is an essential prerequisite for the management of tourist destinations [2]. This study contributes to the literature, by analyzing how tourist tracking methods had been applied. The advantages and disadvantages of each of them were highlighted, suggesting that the GPS is suitable for studies in larger cities/areas and is, in most cases, used simultaneously with questionnaire surveys, allowing for the collection of information related to passenger experiences, such as the attractions visited, the time spent at each attraction, the route taken, the type of transport used, and understanding the levels of satisfaction, the relationships between specific attractions and the intention to revisit or recommend the destiny. On the other hand, Bluetooth is especially suitable for closed places, such as museums, as it allows to understand which works are most visited, and the path taken by the consumer. Smart cards are suitable for public transport, in order to understand how tourists move, the routes they take, and the time it takes for each attraction. Social networks, on the other hand, are associated with the sharing of content by users and interactions in published photographs and videos about a specific attraction or destination.

From the analyzed studies, it is concluded that most of them combine different methods, to fill the gaps of both and to increase the study validity. Although the way this information is available and collected may seem invasive and even scary, these data are very relevant and are used to try to improve experiences for tourists and to assist in the management and planning of services offered.

5 Limitations and Future Research

The most significant limitation of this study is its purely exploratory nature that, without the empirical collection associated, did not allow to obtain more specific findings.

As future research, it would be interesting to try to gather a greater number of studies, to unify the conclusions obtained and reach a consensus on the best applications for each of the available techniques. This knowledge could help future empirical studies, which could be conducted by the Destination Management Organizations themselves, to gather relevant information about the most visited and attractive attractions for tourists, the most used routes, the information that is necessary and where it should be present, and many other aspects that can make a destination more attractive and chosen by visitors.

References

1. Yoshimura, Y., Sobolevsky, S., Ratti, C., et al.: An analysis of visitors' behavior in the Louvre museum: a study using bluetooth data. *Environ Plan B Plan Des* **41**, 1113–1131 (2014). <https://doi.org/10.1068/b130047p>
2. Ferrante, M., De Cantis, S., Shoval, N.: A general framework for collecting and analysing the tracking data of cruise passengers at the destination. *Curr Issues Tour* **21**, 1426–1451 (2018). <https://doi.org/10.1080/13683500.2016.1194813>
3. Bauder, M., Freytag, T.: Visitor mobility in the city and the effects of travel preparation. *Tour Geogr* **17**, 682–700 (2015). <https://doi.org/10.1080/14616688.2015.1053971>
4. Shoval, N., Isaacson, M.: Tracking tourists in the digital age. *Ann. Tour. Res.* **34**, 141–159 (2007). <https://doi.org/10.1016/j.annals.2006.07.007>
5. Martins, M.: O fenómeno backpacker e os seus padrões de movimento espaciotemporal no destino urbano porto. PhD Thesis, Aveiro University (2020). <http://hdl.handle.net/10773/29346>
6. Hartmann, R.: Combining field methods in tourism research. *Ann. Tourism Res.* **15**(1), 88–105. (1988). ISSN 0160-7383. [https://doi.org/10.1016/0160-7383\(88\)90073-4](https://doi.org/10.1016/0160-7383(88)90073-4)
7. Shoval, N., Ahas, R.: The use of tracking technologies in tourism research: the first decade. *Tour. Geogr.* **18**, 587–606 (2016). <https://doi.org/10.1080/14616688.2016.1214977>
8. Yun, H.J., Park, M.H.: Time-space movement of festival visitors in rural areas using a smart phone application. *Asia Pac. J. Tour. Res.* **20**, 1246–1265 (2015). <https://doi.org/10.1080/10941665.2014.976581>
9. Versichele, M., De Groote, L., Bouuaert, M.C., et al.: Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tour. Manag.* **44**, 67–81 (2014)
10. Ali, A., Kim, J., Lee, S.: Travel behavior analysis using smart card data. *KSCE J. Civ. Eng.* **20** (2015). <https://doi.org/10.1007/s12205-015-1694-0>
11. Kovacs, Z., Vida, G., Elekes, Á., Kovalcsik, T.: Combining social media and mobile positioning data in the analysis of tourist flows: a case study from szeged, Hungary. *Sustainability* **13**, 2926 (2021). <https://doi.org/10.3390/su13052926>
12. Barros, C., Moya-Gómez, B., Gutiérrez, J.: Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks. *Curr. Issues Tour.* **23**, 1291–1310 (2020). <https://doi.org/10.1080/13683500.2019.1619674>
13. Caldeira, A.: A experiência de visita dirigida a múltiplas atrações: análise do comportamento espacial do turista e da sua satisfação. PhD Thesis, Aveiro University (2014). <http://hdl.handle.net/10773/12755>
14. Gutiérrez, A., Domènech, A., Zaragoza, B., Miravet, D.: Profiling tourists' use of public transport through smart travel card data. *J. Transp. Geogr.* **88**, 102820. (2020). <https://doi.org/10.1016/j.jtrangeo.2020.102820>
15. Bigné, E., Oltra, E., Andreu, L.: Harnessing stakeholder input on Twitter: a case study of short breaks in Spanish tourist cities. *Tour. Manag.* **71**, 490–503. (2019). <https://doi.org/10.1016/j.tourman.2018.10.013>
16. Rolán, L.X.M., Tymoshchuk, O., Piñeiro-Otero, T., Renó, D.P.: Instagram como red de promoción e hipermediación del turismo rural: el caso de Aldeias Históricas. *Rev. Lat. Comun. Soc.* 1610–1632 (2019)
17. Huertas, A., Martínez-Rolán, X.: Análisis de las fotografías y vídeos de Instagram para la creación de un ranking de popularidad de los territorios y los destinos. *Cuad. Tur.* 197–218 (2020)
18. Casado-Díaz, A.B., Navarro-Ruiz, S., Nicolau, J.L., Ivars-Baidal, J.: Expanding our understanding of cruise visitors' expenditure at destinations: The role of spatial patterns, onshore visit choice and cruise category. *Tour. Manag.* **83**, 104199 (2021). <https://doi.org/10.1016/j.tourman.2020.104199>
19. Spangenberg, T.: Development of a mobile toolkit to support research on human mobility behavior using GPS trajectories. *Inf. Technol. Tour.* **14**, 317–346 (2014). <https://doi.org/10.1007/s40558-014-0005-9>

20. Raun, J., Ahas, R., Tiru, M.: Measuring tourism destinations using mobile tracking data. *Tour. Manag.* **57**, 202–212 (2016). <https://doi.org/10.1016/j.tourman.2016.06.006>
21. Raun, J., Shoval, N., Tiru, M.: Gateways for intra-national tourism flows: measured using two types of tracking technologies. *Int. J. Tour. Cities ahead-of-p* (2020). <https://doi.org/10.1108/IJTC-08-2019-0123>
22. Eccleston, R.E., Hardy, A.L., Hyslop, S.: Unlocking the potential of tracking technology for co-created tourism planning and development: insights from the tourism tracer Tasmania project. *Tour. Plan. Dev.* **17**, 82–95 (2020). <https://doi.org/10.1080/21568316.2019.1683884>
23. Mckercher, B., Hardy, A., Aryal, J.: Using tracking technology to improve marketing: insights from a historic town in Tasmania, Australia. *J. Travel. Tour. Mark.* **36**, 823–834 (2019). <https://doi.org/10.1080/10548408.2019.1580243>
24. Viswanath, S.K., Yuen, C., Ku, X., Liu, X.: Smart tourist-passive mobility tracking through mobile application. In: Giaffreda, R., Cagáňová, D., Li, Y., Riggio, R., Voisard, A. (eds.) *Internet of Things. IoT Infrastructures. IoT360 2014. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 151. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19743-2_28
25. Kádár, B., Gede, M.: Where do tourists go? Visualizing and analysing the spatial distribution of geotagged photography. *Cartogr. Int. J. Geogr. Inf. Geovisualization* **48**, 78–88 (2013). <https://doi.org/10.3138/carto.48.2.1839>
26. Chua, A., Servillo, L., Marcheggiani, E., Moere, A.V.: Mapping Cilento: using geotagged social media data to characterize tourist flows in southern Italy. *Tour. Manag.* **57**, 295–310 (2016). <https://doi.org/10.1016/j.tourman.2016.06.013>
27. Kádár, B.: Measuring tourist activities in cities using geotagged photography. *Tour. Geogr.* **16**, 88–104 (2014). <https://doi.org/10.1080/14616688.2013.868029>
28. Gunter, U., Önder, I.: An exploratory analysis of geotagged photos from instagram for residents of and visitors to Vienna. *J. Hosp. Tour. Res.* **45**, 373–398 (2020). <https://doi.org/10.1177/1096348020963689>
29. Abarca, L.V.E.: Big Data e redes sociais na análise espacial de padrões turísticos em contexto urbano. MsC Thesis, Lisbon University (2016). <http://hdl.handle.net/10451/26022>

Organisation, Classification and Analysis of Online Reviews Directed to Retail in the Municipality of Porto



Pedro Braga, Pedro Quelhas Brito, and Mafalda Teles Roxo

Abstract Web 2.0 has allowed collaboration, interaction and sharing of information online, such as online review platforms. Consequently, these short and straightforward opinions have increasingly proved to be essential sources of information not only for consumers but also for companies, as they represent the consumer's sincere evaluation, free from any kind of bias. In this sense, there should be an interest in the analysis and monitoring of online reviews by companies, as the result of these actions may provide guidelines to readjust their strategy, support decision-making and ensure the satisfaction of their consumers. To generate useful information to assist decision-making and strategies' implementation by retailers in the Municipality of Porto, online reviews from the GoogleMyBusiness platform were organised, classified, and analysed. 9945 online reviews were extracted, directed to 246 retail adaptations of the Municipality of Porto, from 2017 to 2020, which were later classified by the polarity of sentiment (positive, negative, neutral, or mixed). Sentiment analysis was conducted, combined with statistical tests and frequency distribution tables to discover relevant information for retailers. With sentiment analysis, retailers can understand their consumers and their behaviour to adapt their strategies and make the right decisions to ensure their customers' satisfaction. With the results obtained, this study proves that it is possible to extract useful information from online reviews and reveals that it is still an area of little interest for retailers in the Municipality of Porto.

Keywords Municipality of Porto · Sentiment analysis · Online reviews · Retail · Store image · Consumer satisfaction · GoogleMyBusiness

P. Braga · P. Q. Brito · M. T. Roxo (✉)
Faculdade de Economia, Universidade Do Porto, Porto, Portugal
e-mail: mafalda.t.roxo@inesctec.pt

P. Q. Brito
e-mail: pbrito@fep.up.pt

P. Q. Brito · M. T. Roxo
LIAAD/INESC TEC, Porto, Portugal

M. T. Roxo
Universidade da Maia, Maia, Portugal

1 Introduction

The evolution of the Internet has given rise to immense user-generated content and has drastically changed the way we manage, organise, and interact with information [1]. In this evolution, Web 2.0 came to empower users, allowing for collaboration, interaction, and information sharing through online platforms. In this way, it is now possible to express opinions about products or services in a manner visible to everyone through online opinions.

Sentiment analysis is the area that studies and analyses opinions, reviews, and feelings. The constant change in consumer behaviour presents companies with risks and opportunities [2], so companies should develop a monitoring and rapid response resource for several online review platforms [3, 4]. However, existing research on consumer behaviour typically uses questionnaires or other traditional methods [5]. Instead, this research uses sentiment analysis to study the phenomenon independent of manipulation and obtain accurate and natural results and conclusions.

This study aims to organise, classify, and analyse online reviews towards retailers in the Municipality of Porto. The aim is thus to create a set of information to help retailers in the decision-making process. In addition, it seeks to prove the usefulness and ability to extract this information from online comments and understand how retailers in the Municipality of Porto deal with them. The relevance of this study focuses on the importance of understanding the behaviour of consumers by retailers and the desire for continuous improvement to ensure their satisfaction. Furthermore, what consumers say about products or services is an essential issue for companies and associations [6]. Therefore, the exploration of these opinions is of great interest to them and their consumers [7, 8]. Thus, the following research questions were addressed:

1. What is the sentiment and associated intensity present in online reviews written by consumers?
2. How often are store image attributes mentioned, and what is the sentiment associated with each?
3. What is the relationship between sociodemographic variables and the sentiment associated with store image attributes?
4. What is the relationship between store type and the sentiment associated with store image attributes?
5. What is the relationship between sociodemographic variables and rating and intensity?
6. What is the relationship between the rating given by the consumer and the intensity of the feeling present in the reviews?
7. What situations, positive or negative, are frequently highlighted in online reviews by consumers?
8. Do retailers respond to online reviews?

This paper is structured as follows: Chap. 2 introduces the background, concepts, and previous research results; Chap. 3 approaches the proposed methodology. Chapter 4 presents the results; Chap. 5 presents the discussion. Finally, the conclusions and recommendations for future work are drawn.

2 Background

2.1 *The Evolution of Retail and WEB 2.0*

Retail is the commercial activity that consists of the acquisition of products from different branches to provide direct sales of them to consumers in a specific space [9]. Although it is a valid definition, retailers are not limited to selling products, but also a set of services valued by consumers, such as, for example, the availability of facilitated information, the promotion of accessibility, and the creation of an environment, among others [10]. Thus, retail can be classified in several ways, namely the type of products or services it offers.

With the evolution of technology, Web 2.0 emerged, where users interact more with less control [11]. Web 2.0 is the business revolution in the computer industry caused by the shift to the Internet as a platform and an attempt to understand the rules for the success of this new platform [12]. This technological advance has broken down the physical and virtual environment barriers, creating space for the development of omnichannel retailers, where online and offline channels converge to create shopping experiences [13]. Now, consumers switch from channel to channel at their own will [14]. However, the preference for the online or offline channel is very relative and continuously varies depending on a set of factors and consumer needs [15].

2.2 *Store Image*

[16], one of the first investigators in the field describes the store image as the store's personality and how the store is defined in the consumer's mind, partially by its functional qualities and psychological attributes. Later, [17] presented a new, more succinct definition: "store image is the complex perception that a consumer has about the different attributes of a store" (pp. 501). Different authors proposed several store image dimensions in the literature, starting with [16], who found that the store image consisted of layout and architecture, symbols and colours, advertising, and sales personnel. This range of dimensions was developed, namely by [18], and adapted by [19], who presented nine dimensions – sale's point atmosphere, convenience, infrastructure, institutional image, assortment and price level, sales promotions, customer contact employees, and service level - each one related to a set of more specific attributes. The same authors also mention the need for constant monitoring, by the

retailer, of their image in each of the attributes compared to the image of competitors and for making the necessary changes to be preferred when the purchase situation arises.

A study on the influence of attributes on the preference of retailers [20] concluded, through questionnaires, that attributes related to the assortment and price level are the ones that most influence consumers when choosing a retailer. In other words, these types of attributes are the most important for the consumer. Convenience is the next category that influences consumer preferences, followed by an institutional image, customer contact employees, sale's point atmosphere and sales promotions. Another study, conducted by [21], on the same topic but aimed only at female consumers, presented similar results using focus groups. Previous research also suggests that sociodemographic characteristics can influence the perception of store image attributes. In an analysis of different studies on the topic [22–27], the influence of sociodemographic characteristics is clear, but there is a conflict in the results of these studies. As store image is the perception of store attributes, several studies have proven the influence of these attributes on consumer satisfaction [28, 29].

2.3 Online Reviews as a Form of Word of Mouth

The continuous increase in the connection between brands and consumers represents a crucial trend in the times we live in, only possible due to online platforms and devices that allow for faster and more efficient communication [30]. Advances in Web 2.0 allow consumers to share experiences, opinions, and feedback about products, services, or brands through reviews that they can publish online [31]. From this advance, e-WOM arises, defined as "any positive or negative statement made by potential consumers, current or former consumers about a product or company, which is available to all people and institutions through the internet" (pp. 39) [32]. Compared to traditional WOM, e-WOM is more influential thanks to its speed, convenience, reach and lack of face-to-face confrontation [33].

As a form of e-WOM, online reviews express consumer opinions and are published on platforms hosted by the company to which the review is directed or by third parties. These platforms have, in most cases, an open comments section and all reviews are accompanied by a numerical rating [2]. The opinion is defined as what or how an individual thinks about something [34]. In the past, when an organisation needed opinions about its products or services, surveys, focus groups or polls were carried out. However, with the origin of Web 2.0 and social networks, we can currently witness the immense flow of information and opinions present on the Internet that has been helping to restructure businesses and influence consumers [6]. An opinion can be seen as a single sentence, but that single sentence quickly becomes a great source of information. [6] defines opinion as a fivefold (e, a, s, h, t) where e refers to the entity to which the opinion is directed, a refers to the aspect/category of the

entity to which the opinion is directed, s refers to the feeling that can be positive, negative, or neutral, h refers to the author of the opinion, and t refers to the time the opinion was expressed.

Online reviews are gradually becoming important sources of information for consumers [35], as they are considered more credible by consumers compared to company advertising [36, 37] and, therefore, an online review platform is one of the most powerful channels to generate word of mouth [38]. Furthermore, the information present in online reviews can be converted into meaningful knowledge for the company about consumer needs [39]. In this way, online reviews are not only a source of information for the consumer but also for the retailer, as they provide helpful information from the consumer's perspective about the brand, which allows us to understand what can be improved in terms of product or service provision [39].

The constant change in consumer behaviour presents companies with risks and opportunities [2], so companies must develop monitoring and rapid response capabilities on various online review platforms [3, 4]. In addition, retailers must take online reviews into account to design a unique business model and create marketing approaches that will increase loyalty while adding value to consumers [40]. Besides extracting information, retailers must respond to online reviews, as it has emerged as an essential part of managing a brand's reputation [41].

2.4 Sentiment Analysis

Sentiment analysis, also called opinion mining, is the area of investigation that analyses opinions directed at entities to identify consumers' feelings and emotions [6]. In management, this research method is vital to determine opinions about products, services, or brands, understand consumers' attitudes, and improve business practices [6, 42]. There are three types of sentiment analysis, which become more and more specific: (i) at the document level, where the sentiment of an opinion is classified; (ii) at the sentence level, where the feeling of each sentence is classified individually in the same opinion; (iii) at the aspect level, where the feeling of each aspect mentioned in the opinion is classified [6].

The area of consumer behaviour research frequently uses questionnaires, which is a time-consuming method associated with high costs [5]. Furthermore, the quality of data obtained from questionnaires always depends on the willingness of individuals to respond [43]. That said, over the years, it has emerged as a promising alternative to traditional research techniques, as individuals voluntarily give their opinion rather than being required to do so [44].

3 Methodology

A conclusive and descriptive study with a mixed approach was developed to answer the research questions, as it uses quantitative and qualitative research techniques.

3.1 *Data Extraction and Management*

For the intended analysis within the scope of this study, secondary data was collected: online reviews. To decide on the source to extract the data, it was necessary to consider the following criteria: it must be (i) popular and frequently used, (ii) include text comments, (iii) provide sociodemographic data or indications about the authors of the comments. The GoogleMyBusiness platform, included in Google Maps, was the source that best fulfilled all the criteria. A survey of retail establishments located in the Municipality of Porto was carried out through GoogleMaps to guide data extraction, excluding any establishments outside the intended area.

The list of entities prepared in the previous step and the Outscraper software extracted all available reviews, focusing on the time interval: from January 2017 to January 2021. Besides the text of the online reviews, data concerning the reviews, authors, and entities were also collected. Some changes were made to the originally extracted data to organise and clean the sample, such as removing unwanted reviews and assigning gender and origin to authors.

3.2 *Coding and Final Sample*

Within the scope of sentiment analysis, it is necessary to code the reviews in advance according to the sentiment. In 2018, a study was carried out by [45] where the differences in results between the various coding methods were analysed. This study concluded that automatic methods cannot always correctly identify the sentiments present in texts, such as sarcasm and foreign expressions, as was found in the sample of this study. Therefore, it was decided to carry out the coding manually and avoid the use of software indicated for this purpose, to reduce as much as possible the error rate in coding and obtain more reliable analyses.

Once the method was decided, the data was encoded concerning the sentiment and attributes of the entity to which the review was directed (sentiment analysis at document level and aspect level). The sentiment was classified as positive (1), negative (2), neutral (3), and mixed (4). In addition, store image attributes were defined during coding, as they were mentioned in reviews.

To identify the mentioned attributes and the sentiments inherent in them, when a review mentioned an aspect positively, it was coded with 1. When it mentioned an aspect negatively, this aspect was coded with 2. The last stage of the data coding phase

consisted of evaluating the intensity of the sentiment from 1 to 5, where 1 corresponds to Not at all Intense and 5 to Very Intense. Reviews with mixed sentiment were rated in both directions, and reviews with neutral sentiment were not rated.

After collecting the data, the final sample, consisting of 9945 reviews, was drawn from 246 retail establishments, mostly fashion and clothing (86) and food retail (70). In sociodemographic terms, 64.7% of the reviews were written by male individuals and 35.3% by female individuals. Regarding the origin of the authors, 72.6% of the reviews were written by individuals with origins in Portugal, 24.4% by individuals with foreign origins.

4 Results

4.1 Frequency Distribution

Frequency distribution tables were created to analyse the data. Tables 1 and 2 show the distribution of frequencies of Review Sentiment and Intensity of Sentiment. According to the classification, most of the online reviews (76.40%) have a positive sentiment.

Table 1 Frequency distribution of general sentiment

Classification	N	%
Positive	7598	76.40
Negative	899	9.03
Neutral	658	6.62
Mixed	790	7.95

Table 2 Frequency distribution of intensity of sentiment

Classification	M	Rank	N	%
Positive	2.75	1	667	7.95
		2	3055	36.42
		3	2985	35.59
		4	1006	11.99
		5	675	8.05
Negative	2.32	1	269	15.93
		2	890	52.69
		3	307	18.18
		4	163	9.65
		5	60	3.55

Regarding the Intensity of Sentiment, in both positive and negative polarities, the most frequent level of intensity was Little Intense (2) with 36.42% and 52.69%, respectively.

Another frequency distribution table on sentiment directed to store image attributes can be found in Appendix 1, in addition to the associated sentiment, which also shows the incidence value of each attribute in the sample. The attributes General Store Quality, Assortment Variety, Service Quality, Employee Quality and Price, were the most mentioned in the online reviews that make up the sample of this study. All of which have positive sentiment incidence values significantly higher than the values of negative sentiment incidence. However, the attributes Space Dimension, Waiting Times/Queues, Communication with the Store, Post-Purchase Experience, Product Conditions, Stock Management, Hours of Operation, Number of Employees, Capacity, Payment, Problems in the Use of Products and Price Visibility have a negative sentiment incidence value higher than the positive sentiment incidence value. The categories Exceeds or Conforms to Expectations and Innovation have the lowest incidence values, below 0.10%. Finally, the frequency distribution regarding the existence of a response by the entity indicates that only 16.44% of the reviews obtained a response, and 83.56% did not receive any response at all.

4.2 Analysis of Relationships Between Variables Through Association Tests

Chi-square tests were applied to verify whether there is an association between the chosen variables.

Regarding the test applied to the variables Author's Gender and sentiment directed to store image attributes, it was found that the sentiment directed to the attributes Price and Easy Access depends on the Author's Gender.

Regarding the test applied to the Author's Origin and sentiment directed to store image attributes, it was found that the sentiment directed to the attributes Employee Helpfulness, Employee Quality, Product Presentation, Store Location, and Problems in Use of Products depends on the Author's Origin.

Regarding the test applied to the variables Type of Store and sentiment directed to store image attributes, it was found that the sentiment directed to all the attributes depends on the Type of Store.

4.3 Analysis of Relationships Between Variables Through Difference in Means Tests

ANOVA tests were applied to verify differences between the means and conclude if there was any relationship between the variables.

Regarding the test applied to the variable Gender of the Author and the variables Review's Rating, Positive Intensity and Negative Intensity, it was found that the variable Gender of the Author presents differences in the meanings between two or more groups regarding the Positive Intensity only.

Regarding the test applied to the Author's Origin variable and the Review's Rating, Positive Intensity, and Negative Intensity variables, it was found that the Author Origin variable presents differences in the meanings between two or more groups concerning the Review's Rating and Positive Intensity variables.

4.4 Analysis of Correlation Between Variables

The Pearson's Correlation Coefficient (r) test was used to determine the correlation between the variables Review's Rating and Intensity of Sentiment. Two tests were performed for the two types of intensity: positive and negative.

Regarding the test between Review's Rating and Positive Intensity variables, the value of 0.288 indicates a positive correlation between the two variables, although weak. The higher the Positive Intensity of a review, the higher the rating of the review.

Regarding the test between Review's Rating and Negative Intensity, the value of -0.530 indicates a moderate correlation between the two variables. However, it also suggests that the relationship between them is inverse. That is, when the Negative Intensity of a review increases, the rating of the review decreases.

5 Discussion

Based on the analysis of a 9945 online reviews sample, the results are discussed, and research questions are answered. It was found that sentiment towards retailers in the Municipality of Porto is primarily positive. However, negative, neutral, and mixed feelings are still frequent, indicating points for improvement. The intensity associated with the sentiment present in each review is mostly little or reasonably intense, both in negative and positive feelings. However, the average intensity level of positive reviews is higher (RQ1).

Store image attributes are not mentioned with the same frequency, meaning that consumers can value some attributes more or less. The attributes General Store Quality, Assortment Variety, Service Quality, Employee Quality, and Price were the most mentioned in the online reviews that make up the sample of this study,

all of which have positive sentiment incidence values significantly higher than the values of negative sentiment incidence. However, the attributes Space Dimension, Waiting Times/Queues, Communication with the Store, Post-Purchase Experience, Product Conditions, Stock Management, Hours of Operation, Number of Employees, Capacity, Payment, Problems in the Use of Products, and Price Visibility have a negative sentiment incidence value higher than the positive sentiment incidence value. The categories Exceed or Conforms to Expectations and Innovation have the lowest incidence values, below 0.10%. These results do not entirely agree with the results of previous studies mentioned above. However, this analysis focuses on the attributes and not on the dimensions of the store image. In this case, attributes are not analysed as a whole dimension but individually (RQ2).

It was also found that sociodemographic variables can influence the sentiment towards some attributes of the store image. In this study, we found that sentiment directed at Price and Easy Access attributes can depend on the gender of the author. We also found that the sentiment directed at Employee Helpfulness, Employee Quality, Product Presentation, Store Location and Problem in the Use of Products can depend on the author's origin (RQ3). As expected, consumers value different attributes in different types of stores. We found that the sentiment directed to all the detected attributes can depend on the type of store (RQ4).

Regarding the intensity associated with the sentiment, women feel more intensely than men when writing a positive review. Also, foreign consumers give higher ratings and show higher intensities in positive reviews than Portuguese consumers (RQ5). As expected, the intensity of sentiment will affect the rating assigned by the consumer. However, negative intensities have a more significant influence in this regard (RQ6).

In the coding process, it was possible to identify some favourable and unfavourable situations common to certain consumers, namely: some consumers tend to associate the shopping experience with the city or country, most of the neutral reviews pointed to the lack of differentiation, foreign consumers value the efforts that make the experience of buying more natural although they're buying in a different country, and the negligence mentioned by consumers regarding the updating of online information about the store, which results in negative reviews even without the shopping experience occurring (RQ7). Finally, the entity's response rate to online reviews is relatively low in the Municipality of Porto. Also, when there is an answer, it is usually not very personalised (RQ8).

6 Conclusion and Future Work

6.1 Conclusions

The analysis conducted in this paper allowed us to answer the research questions initially proposed, fulfilling the objective of extracting relevant information from online reviews, which were organised and classified to provide an uncomplicated

statistical view of the judgment and sentiment of consumers towards the retail of the Municipality of Porto. This study's main advantage and differentiating factor was the use of secondary data, which allowed us to study the phenomenon itself, excluding any bias that could alter the results. Despite some limitations, such as the time spent and the subjective dimension possibly present, this study proved the feasibility of using and analysing online reviews to create guidelines that help entities make decisions and get to know their consumers better, adopting corrective measures and ensuring their satisfaction. It is expected that the results of this investigation will be used by managers of entities involved in retail in the Municipality of Porto. However, the high lack of attention to monitoring online reviews by them is notorious.

6.2 *Future Work*

In addition to answering several questions, the development of this research created new horizons to be explored. In this sense, this study verifies relationships between variables without presenting associated factors or causes that explain them. In this way, the entities involved could re-adapt their strategies by becoming aware of their causes and taking advantage of or circumventing them for their benefit. Furthermore, although this analysis provides an overview of general information on retail in the Municipality of Porto, it would be interesting to deepen this analysis to a single type of store, making it more specific and targeted. Finally, it would be relevant to observe and compare significant differences in the information obtained by applying the same method to another population, such as a different city or country.

Acknowledgements This work is financed by National Funds through the Portuguese funding agency, FCT—Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

Appendix

See Table 3.

Table 3 Attribute incidence

Attribute	Classification	N	%
General store quality	Positive	2668	26.83
	Negative	35	0.35
	Total	2703	27.18
Assortment variety	Positive	1763	17.73
	Negative	228	2.29
	Total	1991	20.02
Service quality	Positive	1625	16.34
	Negative	217	2.18
	Total	1842	18.52
Employee quality	Positive	1528	15.36
	Negative	200	2.01
	Total	1728	17.38
Price	Positive	1230	12.37
	Negative	400	4.02
	Total	1630	16.39
Available products	Positive	1384	13.92
	Negative	36	0.36
	Total	1420	14.28
Employee helpfulness	Positive	732	7.36
	Negative	96	0.97
	Total	828	8.33
Product quality	Positive	585	5.88
	Negative	46	0.46
	Total	631	6.34
Store atmosphere	Positive	392	3.94
	Negative	30	0.30
	Total	422	4.24
Space dimension	Positive	176	1.77
	Negative	197	1.98
	Total	373	3.75
Store physical aspect	Positive	361	3.63
	Negative	8	0.08
	Total	369	3.71
Store organisation	Positive	193	1.94
	Negative	49	0.49
	Total	242	2.43
Store location	Positive	197	1.98

(continued)

Table 3 (continued)

Attribute	Classification	N	%
	Negative	17	0.17
	Total	214	2.15
Waiting times/queues	Positive	77	0.77
	Negative	136	1.37
	Total	213	2.14
Purchase experience	Positive	116	1.17
	Negative	27	0.27
	Total	143	1.44
Sale's promotions	Positive	114	1.15
	Negative	18	0.18
	Total	132	1.33
Communication with the store	Positive	0	0.00
	Negative	109	1.10
	Total	109	1.10
Post-purchase experience	Positive	23	0.23
	Negative	85	0.85
	Total	108	1.08
Price-quality relation	Positive	92	0.93
	Negative	9	0.09
	Total	101	1.02
Product conditions	Positive	29	0.29
	Negative	45	0.45
	Total	74	0.74
Store cleanliness	Positive	58	0.58
	Negative	13	0.13
	Total	71	0.71
Stock management	Positive	19	0.19
	Negative	49	0.49
	Total	68	0.68
Parking	Positive	38	0.38
	Negative	28	0.28
	Total	66	0.66
Unicity	Positive	60	0.60
	Negative	1	0.01
	Total	61	0.61

(continued)

Table 3 (continued)

Attribute	Classification	N	%
Hours of operation	Positive	27	0.27
	Negative	33	0.33
	Total	60	0.60
Product presentation	Positive	40	0.40
	Negative	9	0.09
	Total	49	0.49
Number of employees	Positive	4	0.04
	Negative	42	0.42
	Total	46	0.46
Capacity	Positive	10	0.10
	Negative	34	0.34
	Total	44	0.44
Stock actualization/news	Positive	41	0.41
	Negative	2	0.02
	Total	43	0.43
Testers and tasters	Positive	32	0.32
	Negative	8	0.08
	Total	40	0.40
Modernised	Positive	31	0.31
	Negative	1	0.01
	Total	32	0.32
Easy access	Positive	22	0.22
	Negative	10	0.10
	Total	32	0.32
Covid-19 measures	Positive	16	0.16
	Negative	15	0.15
	Total	31	0.31
Payment	Positive	4	0.04
	Negative	24	0.24
	Total	28	0.28
Store disposal	Positive	20	0.20
	Negative	8	0.08
	Total	28	0.28
Problems in the use of products	Positive	1	0.01

(continued)

Table 3 (continued)

Attribute	Classification	N	%
	Negative	13	0.13
	Total	14	0.14
Price visibility	Positive	4	0.04
	Negative	6	0.06
	Total	10	0.10
Innovation	Positive	8	0.08
	Negative	1	0.01
	Total	9	0.09
Exceeds or conforms to expectations	Positive	5	0.05
	Negative	4	0.04
	Total	9	0.09

References

1. Tripathi, G., Naganna, S.: Feature selection and classification approach for sentiment analysis. *Mach. Learn. Appl. Int. J.* **2**(2), 1–16 (2015). <https://doi.org/10.5121/mlaij.2015.2201>
2. Hennig-Thurau, T., Walsh, G., Walsh, G.: Electronic word-of-mouth: motives for and consequences of reading customer articulations on the Internet. *Int. J. Electron. Commer.* **8**(2), 51–74 (2003). <https://doi.org/10.1080/10864415.2003.11044293>
3. Becker, K., Nobre, H.: Social network reputation management: an international study. *J. Promot. Manag.* **20**(4), 436–451 (2014). <https://doi.org/10.1080/10496491.2014.930282>
4. Chua, A.Y., Banerjee, S.: Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *J. Am. Soc. Inf. Sci.* **66**(2), 354–362 (2015). <https://doi.org/10.1002/asi.23180>
5. Chrysochou, P.: Consumer behavior research methods. In: *Consumer Perception of Product Risks and Benefits*, pp. 409–428. Springer, Berlin (2017). https://doi.org/10.1007/978-3-319-50530-5_22
6. Liu, B.: *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press (2020)
7. Lin, Y.S., Huang, J.Y.: Internet blogs as a tourism marketing medium: a case study. *J. Bus. Res.* **59**(10–11), 1201–1205 (2006). <https://doi.org/10.1016/j.jbusres.2005.11.005>
8. Schmallegger, D., Carson, D.: Blogs in tourism: changing approaches to information exchange. *J. Vacation Market.* **14**(2), 99–110 (2008). <https://doi.org/10.1177/1356766707087519>
9. Diamond, J., Pintel, G.: *Retailing*. Prentice Hall (1996)
10. Gauri, D.K., Jindal, R.P., Ratchford, B., Fox, E., Bhatnagar, A., Pandey, A., ... Howerton, E.: Evolution of retail formats: past, present, and future. *J. Retail.* **97**(1), 42–61 (2021). <https://doi.org/10.1016/j.jretai.2020.11.002>
11. Aghaei, S., Nematbakhsh, M.A., Farsani, H.K.: Evolution of the world wide web: from WEB 1.0 TO WEB 4.0. *Int. J. Web Seman. Technol.* **3**(1), 1–10 (2012). <https://doi.org/10.5121/ijw est.2012.3101>
12. O'Reilly Archive. <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>. Last accessed 13 May 2021
13. Choudhury, V., Karahanna, E.: The relative advantage of electronic channels: a multidimensional view. *MIS Q.* **32**(1), 179–200 (2008). <https://doi.org/10.2307/25148833>

14. Verhoef, P.C., Neslin, S.A., Vroomen, B.: Multichannel customer management: Understanding the research-shopper phenomenon. *Int. J. Res. Mark.* **24**(2), 129–148 (2007). <https://doi.org/10.1016/j.ijresmar.2006.11.002>
15. Chocarro, R., Cortiñas, M., Villanueva, M.L.: Situational variables in online versus offline channel choice. *Electron. Commer. Res. Appl.* **12**(5), 347–361 (2013). <https://doi.org/10.1016/j.elerap.2013.03.004>
16. Martineau, P.: The personality of the retail store. *Harv. Bus. Rev.* **36**(1), 47–55 (1958)
17. Bloemer, J., De Ruyter, K.: On the relationship between store image, store satisfaction and store loyalty. *Eur. J. Mark.* **32**(5/6), 499–513 (1998). <https://doi.org/10.1108/03090569810216118>
18. Du Preez, R., Visser, E., & Janse Van Noordwyk, H.: Store image: toward a conceptual model Part I. *J. Ind. Psychol.* **34**(2), 50–58 (2008)
19. Pratas, J.M., Brito, P.Q.: *Distribuição: Gestão de pontos de venda e de retalho*, 1st edn. Actual Editora, Lisbon (2019)
20. Thang, D.C.L., Tan, B.L.B.: Linking consumer perception to preference of retail stores: an empirical assessment of the multi-attributes of store image. *J. Retail. Consum. Serv.* **10**(4), 193–200 (2003)
21. Janse van Noordwyk, H. S., Du Preez, R., Visser, E.M.: Importance of apparel store image attributes: perceptions of female consumers. *SA J. Ind. Psychol.* **32**(3), 49–62 (2006)
22. Gehrt, K.C., Yan, R.N.: Situational, consumer, and retailer factors affecting Internet, catalog, and store shopping. *Int. J. Retail Distrib. Manag.* **32**(1), 5–18 (2004). <https://doi.org/10.1108/09590550410515515>
23. Diallo, M.F., Cliquet, G.: Store image perceptions and customer knowledge cues in emerging markets: a cross-country investigation in Brazil and Vietnam. *Int. J. Retail Distrib. Manag.* **44**(12), 1182–1205 (2016). <https://doi.org/10.1108/IJRDM-10-2014-0144>
24. Hyllegard, K., Eckman, M., Descals, A.M., Borja, M.A.G.: Spanish consumers' perceptions of US apparel speciality retailers' products and services. *J. Consum. Behav. Int. Res. Rev.* **4**(5), 345–362 (2005). <https://doi.org/10.1002/cb.18>
25. Moye, L.N., Giddings, V.L.: An examination of the retail approach-avoidance behavior of older apparel consumers. *J. Fashion Market. Manag. Int. J.* **6**(3), 259–276 (2002). <https://doi.org/10.1108/13612020210441355>
26. Oates, B., Shufeldt, L., Vaught, B.: A psychographic study of the elderly and retail store attributes. *J. Consum. Mark.* **13**(6), 14–27 (1996). <https://doi.org/10.1108/07363769610152572>
27. Paulins, V.A., Geistfeld, L.V.: The effect of consumer perceptions of store attributes on apparel store preference. *J. Fashion Market. Manag. Int. J.* **7**(4), 371–385 (2003). <https://doi.org/10.1108/13612020310496967>
28. Doyle, P., Fenwick, I.: How store image affects shopping habits in grocery chains. *J. Retail.* **50**(4), 39–52 (1974)
29. Clotey, T.A., Collier, D.A., & Stodnick, M.: Drivers of customer loyalty in a retail store environment. *J. Serv. Sci. (JSS)* **1**(1), 35–48 (2008). <https://doi.org/10.19030/jss.v1i1.4300>
30. Anastasiei, B., Dospinescu, N.: Electronic word-of-mouth for online retailers: predictors of volume and valence. *Sustainability* **11**(3), 814 (2019). <https://doi.org/10.3390/su11030814>
31. Senecal, S., Nantel, J.: The influence of online product recommendations on consumers' online choices. *J. Retail.* **80**(2), 159–169 (2004). <https://doi.org/10.1016/j.jretai.2004.04.001>
32. Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D.: Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the Internet? *J. Interact. Mark.* **18**(1), 38–52 (2004). <https://doi.org/10.1002/dir.10073>
33. Phelps, J.E., Lewis, R., Mobilio, L., Perry, D., Raman, N.: Viral marketing or electronic word-of-mouth advertising: examining consumer responses and motivations to pass along email. *J. Advert. Res.* **44**(4), 333–348 (2004)
34. Oxford English Dictionary. <https://oed.com/>. Last accessed 25 Apr 2021
35. Mosteller J., Mathwick C.: Consumers helping consumers: The role of psychological need fulfillment in an online reviewer community. In: Obal M., Krey N., Bushardt C. (eds) *Let's*

- Get Engaged! Crossing the Threshold of Marketing's Engagement Era, pp. 377–378. Springer, Berlin (2016). https://doi.org/10.1007/978-3-319-11815-4_106
36. Bickart, B., Schindler, R.M.: Internet forums as influential sources of consumer information. *J. Interact. Mark.* **15**(3), 31–40 (2001). <https://doi.org/10.1002/dir.1014>
 37. Godes, D., Mayzlin, D.: Using online conversations to study word-of-mouth communication. *Mark. Sci.* **23**(4), 545–560 (2004). <https://doi.org/10.1287/mksc.1040.0071>
 38. Dellarocas, C.: The digitisation of word of mouth: promise and challenges of online feedback mechanisms. *Manage. Sci.* **49**(10), 1407–1424 (2003). <https://doi.org/10.1287/mnsc.49.10.1407.17308>
 39. Ibrahim, N.F., Wang, X.: Decoding the sentiment dynamics of online retailing customers: Time series analysis of social media. *Comput. Hum. Behav.* **96**, 32–45 (2019). <https://doi.org/10.1016/j.chb.2019.02.004>
 40. Ramanathan, U., Subramanian, N., Parrott, G.: Role of social media in retail network operations and marketing to enhance customer satisfaction. *Int. J. Oper. Prod. Manag.* **37**(1), 105–123 (2017). [10.1108/IJOPM-03-2015-0153](https://doi.org/10.1108/IJOPM-03-2015-0153)
 41. Sparks, B.A., Bradley, G.L.: A “Triple A” typology of responding to negative consumer-generated online reviews. *J. Hospitality Tourism Res.* **41**(6), 719–745. [10.1177/1096348014538052](https://doi.org/10.1177/1096348014538052)
 42. Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., By, T.: Sentiment analysis on social media. In: 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 919–926. IEEE (2012). <https://doi.org/10.1109/ASONAM.2012.164>
 43. Groves, R.M.: Nonresponse rates and nonresponse bias in household surveys. *Public Opin. Q.* **70**(5), 646–675 (2006). <https://doi.org/10.1093/poq/nfl033>
 44. Decker, R., Trusov, M.: Estimating aggregate consumer preferences from online product reviews. *Int. J. Res. Mark.* **27**(4), 293–307 (2010). <https://doi.org/10.1016/j.ijresmar.2010.09.001>
 45. Chapman, L., Resch, B., Sadler, J., Zimmer, S., Roberts, H., Petutschnig, A.: Investigating the emotional responses of individuals to urban green space using twitter data: a critical comparison of three different methods of sentiment analysis. *Urban Plan.* **3**(1), 21–33 (2018). <https://doi.org/10.17645/up.v3i1.1231>

Strategic Digitization of Tourism



José J. Galán, Ramón A. Carrasco, and Antonio Latorre

Abstract The tourism sector is undergoing an accelerated digital transformation, augmented by the current pandemic, and tourists must adapt to this new environment. There are many options for digitalisation in the tourism sector and their success depends on the grade of tourist satisfaction. Not all options work and, from a marketing point of view, it is important to know this grade of satisfaction in order to be able to offer the technological experience that the tourist expects. In this context, assessing which technological option is more important than another is a complex decision problem. In this type of problem, the AHP model, an analytical and hierarchical decision-making process, has been used successfully and is extensively recognised. The goal of this paper is to offer an adaptation of the AHP model on the TAM, a model that measures the grade of user satisfaction of technology, to find out the grade of tourist satisfaction towards the technological alternatives used in tourism.

Keywords Tourist · Tourism · Tam · Ahp · Smart tourism · Strategic digitization

1 Introduction

Tourism is a sector that has been evolving technologically in a sequential manner, and with the Covid-19 pandemic it was necessary to increase this momentum in order to

J. J. Galán (✉)

Dpto. de S. Informáticos y Computación, Facultad de Estudios Estadísticos, Madrid, Spain

e-mail: josejgal@ucm.es

R. A. Carrasco

Dpto. de Organización de Empresas Y Marketing, Facultad de Comercio Y Turismo, Madrid, Spain

A. Latorre

CCS, Universidad Politécnica de Madrid, 28660 Madrid, España

J. J. Galán · R. A. Carrasco

Universidad. Complutense de Madrid, Madrid, España

minimise as much as possible the damage suffered in this sector by avoiding human contact [1].

But are all potential tourists prepared for this technological adaptation, what are the key technological elements and how do they affect the potential tourist? This paper sets out a strategy to determine which factors and to what grade affect the digitisation of the future technological tourist.

In point 2, this work carries out a study of the work already carried out and found in the scientific database Scopus on the application of a methodology to measure technological acceptance in tourism. We see that there are no relevant articles on the subject, presenting in point 3 the methodology that complements this: we will use the TAM model, which evaluates the grade of acceptance by a user of the use of technology and the AHP methodology so that with this decision-making process we can determine through the steps described in point 4 the grade of importance that the different technological elements have for a tourist to accept the use of a certain technology.

2 Related Work

In the Scopus database, which is widely accepted as a scientific source [2], the references described in Table 1 to the TAM and AHP model in relation to Tourism have been found as of 17 September 2021.

In the related works that we have found, there is no reference to a model that resolves how to know the grade of satisfaction of a tourist in the use of technology. Therefore, applied to tourism, we decided to solve this lack by applying the AHP model to TAM, presented in point 3 and developed in point 4.

3 Methodology Used

In order to assess the grade of tourist acceptance of the technology, we will use the technology acceptance model (TAM) [3], created by Fred D. Davis in 1989 and since then considered the basis for the study of user acceptance of a technological change [4]. Several versions have been made since then and adaptations for different business models and sectors have been made [5–7].

To feed TAM it is necessary to decide which technology is the most appropriate based on certain criteria that we will describe in Point 4. For this decision we will use the analytical hierarchy process (AHP) [8]. Thomas Saaty developed this hierarchical method, which makes use of qualitative and quantitative analysis to help make decisions. It consists of ranking and choosing alternatives to solve unstructured multi-criteria decision-making problems.

Table 1 Work related to TAM and AHP in tourism

Title	Author	Year	Source	Thematic	Tourist area
Digital tourist: Variables that define their purchasing behaviour [Turista digital: variables que definen su comportamiento de compra]	Miranda, D.D., Briley, D	2021	Tourism research (21), pp. 1–21	Acceptance of robots in hotels after the pandemic	Yes
The consequences to browse in spanish vs. english when developing loyalty towards a tourist destination on internet [Las consecuencias de navegar en español vs. inglés a la hora de desarrollar la lealtad hacia un destino turístico en internet]	Alcántara-Pilar, J.M., Barrio-García, S.D	2017	Management notebooks 17(1), pp. 57–82	Analyses the influence of the cultural values of a language when processing information from a website to a tourist destination	Yes

4 Strategic Digitization of Tourism

For the purpose of this study, we will apply the AHP technique, making use of RStudio, Microsoft Excel and Microsoft Visio, performing the 4 steps shown in Fig. 1.

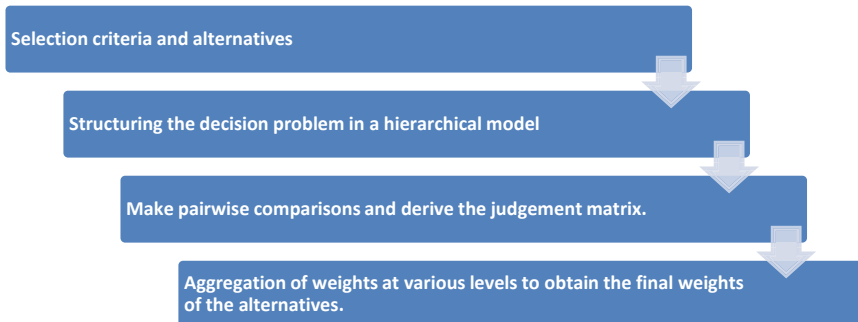


Fig. 1 Steps of the proposed digitisation strategy

4.1 Step 1: Selection Criteria and Alternatives

By studying scientific articles and reviews with a high impact index, we have considered the criteria in Fig. 2 and the alternatives in Fig. 3, both of which are described below.

- **Technological factors:** technology is necessary in tourism [9], the easier it is to use, and the more tourists will accept its use.
- **Resistance to change:** we all have a resistance to get out of our comfort zone [10] and adopt new ways [11].

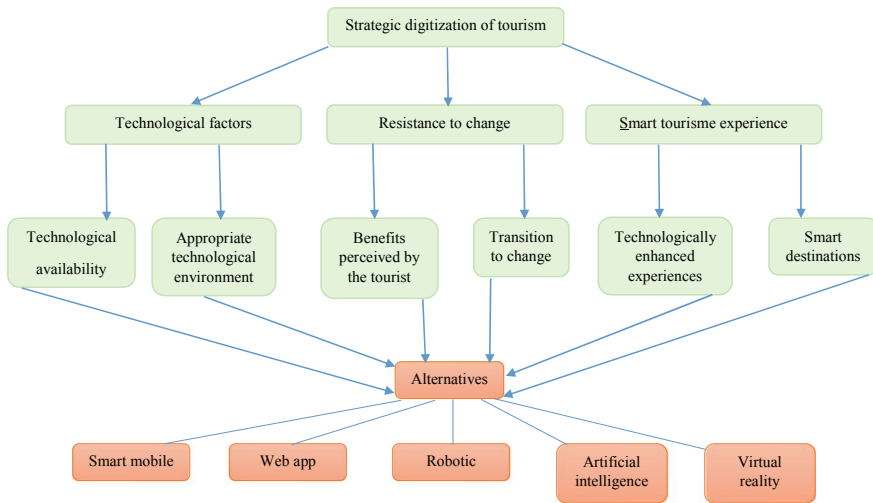


Fig. 2 Strategic digitalization of tourist value tree

	Weight	Robotic	Smart mobile	Artificial intelligence	Web app	Virtual reality	Inconsistency
Strategic digitization of tourism	100.0%	26.3%	24.3%	22.5%	15.7%	11.1%	0.5%
Resistance to change	59.5%	15.7%	11.9%	16.5%	9.0%	6.3%	0.0%
Benefits perceived by the tourist	44.7%	10.9%	8.3%	14.4%	6.3%	4.8%	4.4%
Transition to change	14.9%	4.8%	3.6%	2.1%	2.8%	1.6%	4.4%
Smart tourism experience	27.6%	8.0%	8.4%	4.5%	3.5%	3.3%	0.0%
Technologically enhanced experiences	16.6%	4.4%	5.7%	3.0%	1.4%	2.1%	1.9%
Smart destinations	11.1%	3.6%	2.7%	1.6%	2.0%	1.2%	4.4%
Technological factors	12.8%	2.5%	4.0%	1.5%	3.2%	1.5%	0.0%
Technological availability	7.7%	1.5%	2.6%	1.0%	1.9%	0.8%	4.7%
Appropriate technological environment	5.1%	1.1%	1.5%	0.5%	1.4%	0.7%	2.3%

Fig. 3 Summary of the final weightings of the alternatives

- **Smart tourism experience:** facilitating management [12], automating traditional services and facilitating the tourist experience [13].
- **Technological availability:** having the necessary technology [14].
- **Appropriate technological environment:** the means used must be able to be used in a suitable environment [15] and meet the minimum requirements.
- **Benefits perceived by the tourist:** the user must perceive that his or her experience is enriched by the use of these technologies [16].
- **Transition to change:** using technology instead of traditional methods [17].
- **Technologically enhanced experiences:** all the means employed contribute to enhance and enrich the experience [18].
- **Smart mobile:** through apps, users interact with destinations [19], routes, travel agencies, museums, etc.
- **Web app:** allows a trip to be managed from a device with internet. [20].
- **Robotic:** automating traditionally human processes [21].
- **Artificial intelligence:** by accessing big data and machine learning to establish routes [22] and preferences according to each customer.
- **Virtual reality:** Visualise scenarios that no longer exist [23], replace guides by avoiding contact, etc. The different levels are in Tables 2, 3 and 4.

4.2 Step 2: Structuring the Decision Problem into a Hierarchical Model

To facilitate understanding, a visual and hierarchical model is constructed with inter-related criteria, decomposing the decision problem into elements according to their common characteristics [24]. The proposed model (Fig. 2) is a hierarchical structure of 4 levels of multi-criteria decision, according to their dimensions and sub-dimensions, ensuring that they are different and do not overlap and guaranteeing coherence.

The first level is the objective, the digitisation of tourism. The second level shows 3 major dimensions that influence the TAM adapted to tourism as we have seen in point 3. The third level is the criteria that are widely accepted for their influence on the factors of level 2. It should be noted that no decision alternatives have been considered and that all the levels described above contribute to the goal.

4.3 Step 3: Make Pairwise Comparisons and Derive the Judgement Matrix

For each element “e” (objective) of a hierarchy level, the elements (dimensions) of the next level down in the hierarchy are compared in pairs with respect to their influence on the element “e”. The importance of each dimension over another and in relation to the objective is represented through the absolute number rating scale (1–9)

Table 2 Elements of the two-dimension tourist technology alternatives model, level 2

Criteria/alternatives	Technological availability	Appropriate technological environment	Benefits perceived by the tourist	Transition to change	Technologically enhanced experiences	Smart destinations
Technological factors	It is essential to have the necessary technology: smartphone, computer, etc.	In addition to having technology, this must be adequate (e.g., a mobile phone must have smart technology and internet access)	Seeing technology as an ally, not a problem	Interest in using technology instead of traditional media	See how technology improves the quality of your experience	Smart destinations are a consequence of using technology
Resistance to change	Having the necessary technology in place makes it easier for them to be accepted	The easier the technology we have, the easier it will be to accept it	If we perceive benefit then there will be less resistance to its use	Less problems in the adaptation process will mean less resistance	By demonstrating that technology improves expediency, resistance is reduced	In a smart destination, resistance disappears because of the obligation to adapt
Smart tourism experience	Technology is necessary in smart tourism	Have the technology required specifically to improve the experience (e.g. use recommended browser)	The facilities and conveniences provided by automated systems should be appreciated	Use the improvements offered	Directly related, both improve together	The smart destination governs smart tourist experience

Table 3 Elements of the two-dimension tourist technology alternatives model, level 3

Criteria/alternatives	S1: smart mobile	S2: web app	S3: robotic	S4: artificial intelligence	S5: virtual reality
C1: technological availability	Training on how to use a smartphone	Training on how to use a website (e.g. shopping online)	Use mechanisms that automate the experience (e.g. smart hotels)	Leveraging suggested routes based on machine learning	Learning through virtual reality (e.g. museum assistant)
C2: appropriate technological environment	Availability of a smartphone	Availability of access to a web browser	Automations must be accessible to all publics	The format in which the information is received (e.g. mobile phone, brochure, etc.)	In addition to the technology, the space, light and sound must be appropriate
C3: benefits perceived by the tourist	Value the mobile as an enhancement to the experience (e.g., apps for sharing photos and comments)	Understand the agility of its use (e.g. shopping without leaving home and contactless)	Automate what does not add value	Through the study of data, the experience is improved	It is possible to represent with images what could only be told (e.g. showing dinosaurs in a museum)
C4: transition to change	Use apps in a natural way (e.g. consult an app instead of a guide)	Be willing to engage digitally. (For example, requesting information through a form)	Accept to be served by an automated system instead of a person	Use social networks and digital environments	Relying on digital media. (E.g. accept the explanation of a hologram)
C5: technologically enhanced experiences	It is possible to represent with images what could only be told (e.g. showing dinosaurs in a museum)	Knowing how to use the web as a complementary experience	Complete the tourist experience with the facilities provided by automated mechanisms	Apps, web and elements adapted to the user's interests	It complements what could not be seen by normal means
C6: smart destinations	Destinations that require the use of smartphones to use their resources	Use of resources at destination via the web	Smart destinations use robots in their interactions	User and their interests, adapting the destination to their preferences	Showing topics of interest according to user preferences

Table 4 The fundamental scale of absolute numbers

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgement slightly favour one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgement strongly favour one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favouring one activity over another is of the highest possible order of affirmation
Reciprocals of above		If activity <i>i</i> has one of the above non-zero numbers assigned to it when compared with activity <i>j</i> , then <i>j</i> has the reciprocal value when compared with <i>i</i>
1.1–1.9	If the activities are very close	May be difficult to assign the best value but when compared with other contrasting activities the size of the small numbers would not be too noticeable, yet they can still indicate the relative importance of the activities

Source Saaty [27]

in Table 5, where a higher rating is considered superior compared to the other [25]. The values assigned to each element of the matrix must meet 3 rules to be consistent, these rules are $a_{ij} > 0$; $a_{ij} = 1/a_{ji}$ (reciprocity) and $a_{ii} = 1$ for all *i*. This measure is called the consistency scale (CR) and is considered consistent and acceptable when it is less than 0.10. [26]

For the problem shown in Fig. 2, we have obtained three dimensions that are compared in pairs (PCM) with respect to the objective. Then, for each of these dimensions, a calculation is made of their weights in pairwise comparison with their sub-dimensions. Finally, each of these sub-dimensions has 5 dimensions with a direct relationship (alternatives) on which the weight is calculated in relation to the higher dimension.

4.4 Step 4: Local Weights and Consistency of Comparisons to Obtain the Final Weights of the Alternatives

Next, we calculate the local weights using the eigenvector method (EVM) [27] by Eq. (1), where A represents the comparison matrix [25], w the eigenvector or preference vector, and λ_{\max} the eigenvalue, for which 2 theorems must be satisfied: $\lambda_{\max} \geq n$; and A is consistent if and only if $\lambda_{\max} = n$. Where λ_{\max} is the maximum eigenvalue and n is the dimension of the decision matrix [8].

$$Aw = \lambda_{\max}w \quad (1)$$

To ensure the credibility of the single hierarchical ordering, it is necessary to test the consistency of the hierarchical structure by calculating the random consistency ratio. The ratio to estimate the consistency is the Consistency Ratio (CR), represented in (2), and is obtained by dividing the Consistency Index (CI) which takes the value $(\lambda_{\max} - n)/(n - 1)$, and the Random Consistency Index (RI) represents the consistency of a randomly generated PCM [25].

$$CR = CI/RI \quad (2)$$

Only when $CR \leq 0.10$ per PCM is met will there be coherence, otherwise it will be necessary to adjust the values of the PCM elements [26]. The results of this analysis are shown in Fig. 3.

Finally, the weight of the overall priority vector is calculated, and the importance of the weights of the dimensions and sub-dimensions is also determined, using the multiplicative AHP approach [28]. By definition, the weights of the alternatives and the importance of the criteria are normalised to sum to unity [26]. With this information (Fig. 3.) the marketing department can make decisions to satisfy its objective. It can be seen that the Resistance to change segment prevails (59.5%) and that Robotic (26.3%) and Smart mobile (24.3%) are the most important alternatives.

5 Conclusions

There is a clear resistance to the use of technology, therefore marketing departments have a great task ahead to make tourists aware of the benefits (44.7% importance) of technology to improve their experience. Mainly promoting the use and training in the use of automated processes (26.3%) and in the use of Smart mobile (24.3%).

The concern and need to have technological elements is lower, evidently, access to technology and means are common and common among those who have the possibility of accessing tourism.

Artificial intelligence (22.5%) determines preferences and regulates customer experience and satisfaction, and is a vital source of information for marketing.

References

1. Zeng, Z., Chen, P.-J., Lew, A.A.: From high-touch to high-tech: COVID-19 drives robotics adoption. <https://doi.org/10.1080/14616688.2020>
2. Moreno, C., González, R.A.C., Viedma, E.H.: Data and artificial intelligence strategy: a conceptual enterprise big data cloud architecture to enable market-oriented organisations. *IJIMAI* **5**(6), 7–14 (2019)
3. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **13**(3), 319–340 (1989)
4. Al-Nuaimi, M.N., Al-Emran, M.: Learning management systems and technology acceptance models: A systematic review. *Educ. Inf. Technol.* **26**(5), 5499–5533 (2021). <https://doi.org/10.1007/s10639-021-10513-3>
5. Casey, J.E., Pennington, L.K., Mireles, S.V.: Technology acceptance model: assessing preservice teachers' acceptance of floor-robots as a useful pedagogical tool. *Technol. Knowl. Learn.* **26**(3), 499–514 (2021). <https://doi.org/10.1007/s10758-020-09452-8>
6. Khoza, S., de Beer, L.T., van Niekerk, D., Nemaikonde, L.: A gender-differentiated analysis of climate-smart agriculture adoption by smallholder farmers: application of the extended technology acceptance model. *Gend. Technol. Dev.* **25**(1), 1–21 (2021). <https://doi.org/10.1080/09718524.2020.1830338>
7. Sek, Y., Lau, S., Teoh, K., Law, C., Parumo, S.B.: Prediction of user acceptance and adoption of smart phone for learning with technology acceptance model. *J. Appl. Sci.* **10**(20), 2395–2402 (2010). <https://doi.org/10.3923/jas.2010.2395.2402>
8. Saaty, T.L.: How to make a decision: the analytic hierarchy process. *Eur. J. Oper. Res.* **48**(1), 9–26 (1990)
9. Bagherinia, A., Zali, N.: Urban tourism development mechanisms based on information technology. [Desenvolvimento do turismo urbano baseado na tecnologia da informação] *Anuario do Instituto De Geocienc.* **36**(1), 92–97 (2013). https://doi.org/10.11137/2013_1_92_97
10. Krügel, J.P., Traub, S.: Reciprocity and resistance to change: an experimental study. *J. Econ. Behav. Organ.* **147**, 95–114 (2018). <https://doi.org/10.1016/j.jebo.2017.12.017>
11. Kiran, D.R.: Total quality management: key concepts and case studies. Butterworth-Heinemann (2016)
12. Hailey Shin, H., Jeong, M., Cho, M.: The impact of smart tourism technology and domestic travelers' technology readiness on their satisfaction and behavioral intention: a cross-country comparison. *Int. J. Tourism Res.* **23**(5), 726–742 (2021). <https://doi.org/10.1002/jtr.2437>
13. Vizueté, D.D.C., Montoya, A.V.G., Jácome, E.A.M., Velásquez, C.R., Borz, S.A.: An evaluation of the importance of smart tourism tools in the riobamba canton, ecuador. *Sustain. (Switzerland)* **13**(16) (2021). <https://doi.org/10.3390/su13169436>
14. Nabi, M., Toeroe, M., Khendek, F.: Availability in the cloud: state of the art. *J. Netw. Comput. Appl.* **60**, 54–67 (2016)
15. Ko, W.L., Kim, S.Y., Lee, J., Song, T.H.: The effects of strategic alliance emphasis and marketing efficiency on firm value under different technological environments. *J. Bus. Res.* **120**, 453–461. (2020) <https://doi.org/10.1016/j.jbusres>
16. Madero Gómez, S.M., Barboza, G.A.: Interrelación de la cultura, flexibilidad laboral, alineación estratégica, innovación y rendimiento empresarial. *Contaduría y administración*, **60**(4), 735–756 (2015). <https://doi.org/10.1016/j.cya.2014.08.001>
17. Jahanmir, S.F., Cavadas, J.: Factors affecting late adoption of digital innovations. *J. Bus. Res.* **88**, 337–343 (2018). <https://doi.org/10.1016/j.jbusres.2018.01.058>
18. Temel Eginli, A., Isik, S.: Generational differences in digital age a research on technology experiences of generations. *Int. J. Sci. Technol. Res.* **9**(2), 3150–3154 (2020)
19. Chen, P.: Recommendation strategies for smart tourism scenic spots based on smart city (2021). https://doi.org/10.1007/978-3-030-62746-1_108
20. Aggarwal, S., Gour, A.: Peeking inside the minds of tourists using a novel web analytics approach. *J. Hospitality Tourism Manag.* **45**, 580–591 (2020). <https://doi.org/10.1016/j.jhtm>

21. Merkurjeva, G., Bolshakov, V., Martins, J., Gonçalves, R., Benito, G.B.: Intelligent simulation for tourism and cultural heritage management. In: Paper presented at the 28th European Modeling and Simulation Symposium, EMSS, pp 328–333 (2016)
22. Xia, H., An, W., Liu, G., Hu, R., Zhang, J.Z., Wang, Y.: Smart recommendation for tourist hotels based on multidimensional information: a deep neural network model. *Enterp. Inf. Syst.* 1–25 (2021). <https://doi.org/10.1080/17517575.2021.1959651>
23. Alyahya, M., McLean, G.: Examining tourism consumers' attitudes and the role of sensory information in virtual reality experiences of a tourist destination. *J. Travel Res.* 00472875211037745 (2021). <https://doi.org/10.1177/00472875211037745>
24. Cock, G., Rendon, J.F.L.: Ludica aplicada al problema de localizacion de planta fisica utilizando analytic hierarchy process (ahp), y solver para excel. *Scientia et technica* **2**(42) (2009). <https://doi.org/10.22517/23447214.2609>
25. Ocampo, L., Alinsub, J., Casul, R.A., Enquig, G., Luar, M., Panuncillon, N., Ocampo, C.O.: Public service quality evaluation with SERVQUAL and AHP-TOPSIS: A case of Philippine government agencies. *Socio-Econ. Plan. Sci.* **68**, 100604 (2019). <https://doi.org/10.1016/j.seps.2017.12.002>
26. Subramanian, N., Ramanathan, R.: A review of applications of analytic hierarchy process in operations management. *Int. J. Prod. Econ.* **138**(2), 215–241 (2012)
27. Saaty, T.L.: Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* **1**(1), 83–97 (2008)
28. Saaty, T.L.: *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. McGraw-Hill, New York (1980)

Tourism Professionals' Awareness of the New Technologies: A Premise of the Destinations' Modernization and Development



Diana Foris , Gabriela-Ramona Bâra, and Simona Mălăescu 

Abstract In the context of the current requirements regarding the modernization and the development of tourist destinations in order to remain competitive and attractive for the modern tourist, the use of new technologies becomes an urgent necessity, representing a reality of the present and especially of the future. People's attitudes on the use of emergent technologies in tourism varies from stakeholders to practitioners and tourists. In this respect, a mix of qualitative marketing research based on semi-structured interviews was conducted with practitioners. The results of the study highlighted on the long-term positive attitude of tourism practitioners towards the use of new technologies in tourism, especially during pandemic and the positive effects on the development and modernization of tourist destinations. On the short-term, an ambivalent reaction probably due to the technological and digital “binge” during pandemic was registered. The practical implications of the results for the stakeholders imply that the abrupt and salient use of digital services and new technologies could provoke on the short term an ambivalent response from users. The paper also offers suggestions for future research directions.

Keywords Tourist destination · New technologies · IoT · VR · AI · Web 4.0

1 Introduction

The development and progress of new technologies, including information and communication technology, have had a strong impact on tourism and the hospitality

D. Foris (✉)

Faculty of Food and Tourism, Transilvania University of Brasov, Brasov, Romania
e-mail: diana.foris@unitbv.ro

Development and Innovation, CiTUR—Centre for Tourism Research, Peniche, Portugal

G.-R. Bâra

Faculty of Food and Tourism, Transilvania University of Brasov, Brasov, Romania

S. Mălăescu

Human Geography and Tourism Department, Babes-Bolyai University, Cluj-Napoca, Romania
e-mail: simona.malaescu@ubbcluj.ro

industry, and recently the use of modern technology in tourism has been expanded in order to offer a wider range of products and services to tourists from the whole world. Regarding the implications of technology on tourism, it can be stated that “information is the lifeblood of tourism” [1] and information and communication technology (ICTs) is the mainstay of the concept of everything that is smart, including smart tourism [2]. With the involvement of ICT in the rapid evolution of tourism, the innovation activity in tourism has increased substantially in intensity [3]. The widespread adoption of new technologies in the tourism and hospitality industry in recent years has fundamentally reshaped the way services are provided and received [4]. Therefore, it can be said that tourism is constantly undergoing radical changes and, at the same time, rapid, in terms of technology. Recent technological development has led to an unprecedented revolution in the hospitality and tourism industry [5]. To improve the quality of services and the customer experience, the tourism and hospitality industry has widely adopted information technology (IT) [6]. The tourism industry can be considered as a critical field of testing and application for informatics, which, in turn, allows new services, products, and cooperation for the field of tourism [7]. Therefore, understanding how tourism experiences can be improved through technology is very important [8].

The purpose of this paper is to identify the level of awareness and the perception about the use of technology in the tourism sector as a premise for the development and modernization of a tourist destination, on a Romanian context.

2 Literature Review

Technological innovations bring together the full range of stakeholders in the ecosystems of tourism services, and tourism experiences with technological involvement offer tourists the opportunity to create value at all stages of the journey [9]. New digital technologies (IoT, VR, AR, AI, connectivity, Web 4.0) represent the technical platform that allows the convergence of the concepts of Tourism 4.0 and smart tourism, with a content that overlaps in the context of the fourth industrial revolution [10]. One of the widest applications in tourism that new technologies register is that of improving the visitor experience to the cultural objectives of historical-cultural heritage and increasing the quality of heritage interpretation [11].

The Internet of Things (IoT) technology has a multiple impacts on the tourism industry, along with the requirement for mobility and flexibility throughout the travel and accommodation experience [12]. The adoption of IoT technologies redesigns and improves the performance of all the main processes in the field of hospitality and in the tourism sector [13]. The system of the Internet of Things implies that every Thing communicates, every Thing is detectable, every Thing interfaces [10]. IoT technology is the future, in which anything can be adjusted, controlled or verified

using technology. In the tourism industry, IoT technology involves the use of interconnected devices equipped with sensors, software and that allow remote communication, and information processing such as geolocation bracelets, mobile guide apps, ticket payment apps, virtual reality technologies, etc. [14].

Virtual Reality (VR) is a rapidly developing technology with a wide range of uses, and the tourism industry, especially destinations, could benefit from its potential [15]. The Virtual Reality (VR) can help tourists to experience a new digital environment, different from the current physical environment, creating a spatial presence in a certain tourist destination [16] and can answer questions in the process pre-purchase helping potential tourists decide [17]. In using VR in marketing strategy, the destinations have a substantial influence on the marketing process, on information and on the decision-making process within potential tourists. Tourists can experience in advance the destination, atmosphere and possible emotions [15], and VR technology could be used in tourism to allow an experience of the product before purchase and in providing emotional experience [18].

Augmented Reality (AR) refers to the improvement of the real environment through computer-generated content, which is updated and supplemented with graphic content [19]. Augmented Reality offers many opportunities for mobile applications, which can be exploited in industries such as tourism, and, due to the context of the immediate location, the use of augmented reality devices has the potential to create the next generation of computerized travel guides [20]. Moreover, augmented reality contributes to the enrichment of the tourist experience especially in visiting cultural objectives of appreciable size whose interior design, for various reasons, do not have much to offer tourists in terms of designed interpretation [21]. Both VR and AR meet much greater demand for active involvement during the visit of a tourist destination by tourists from the younger generations of tourists that are more oriented to technology [22].

Technologies based on artificial intelligence (AI) have redefined the hospitality industry [23]. Artificial intelligence technology has become an important engine for the development of modern tourism and has penetrated every link in the tourism industry [24]. AI offers automated, personalized and intelligent travel services, allows tourists to learn about their behaviors and offers a personalized experience. Technological progress determines the need to convert all tourism resources into smart resources, with the help of sensors, Big Data, new ways of connectivity and data exchange [25]. Therefore, innovation and technology can contribute to sustainable growth of the tourism [26]. New emerging technologies such as chatbots, virtual reality, language translators, etc. can be effectively applied in the tourism and hospitality industry [27], as well as at the level of tourist destinations.

The use of modern technologies becomes a necessity of many tourist destinations in order to remain competitive and attractive for the modern tourist [28].

The tourist destination is a complex system consisting of a variety of organizations that offer tourist products and services (accommodation, transport, attractions, etc.), natural and anthropic resources, as well as other units that provide assistance services for tourists [29]. According to systems theory, the destination can be seen as a system in which the different component subsystems are interconnected through material

flows/exchanges, information flows and energy flows [30]. Therefore, the destination is a system, a network of connections, and the management of the destination plays an essential role in its development. In this context, the new technologies are able to provide tools and solutions for a good management of the situations generated by the fact that the development of tourist activities at the destination level takes place in a competitive and constantly changing environment, as well as the possibility that the destination offers valuable experiences for tourists.

The rapid development of technologies at the level of a tourist destination involves the use of intelligent tools and solutions at the level of all organizations in that destination as well as at the community level. Thus, appears the concept of Smart Travel Destinations (STD)—a concept that comes from the development of smart cities—meaning that, by incorporating technology in all organizations and entities, destinations will exploit synergies between ubiquitous detection technology and their social components to support the enrichment of tourism experiences [31]. Moreover, Buhalis and Amaranggana identified the models and types of VR and AR tourism applications in Smart Tourist Destinations that help to digital experimentation of the environment of tourist sites; tracking systems and monitoring of vehicles for real-time information and vehicle locations, energy-efficient green hotels to reduce electricity consumption and pollution, applications for displaying available travel packages, translation and providing guidance to tourists during the visit, access to inform keep in mind about nearby attractions to visit and visiting hours via mobile devices [31].

By introducing IoT innovation, smart and sustainable cities can improve various aspects of their urban management, for example, urban mobility, public transport, e-government, safety, security, public lighting and environmental monitoring [32]. Moreover, the adoption of IoT technologies is expected to enable the monitoring, control and management of all available resources. Therefore, IoT technology can also be applied to smart tourist destinations.

The Artificial Intelligence (AI) studies and develops theories, methods, technologies and applications to simulate and expand human intelligence; as a branch of computer science, AI is the intelligence demonstrated by machines and includes robotics, language recognition, image recognition, natural language processing [33]. AI technology has become an important engine for modern tourism development, it has penetrated every link in the tourism industry and it plays an important role in promoting the development of smart tourism—being conducive to the collection and research for tourist information by tourists and the dissemination of tourist information by tourism operators [33]. In this sense, AI plays an important role in improving the quality of destination management, in increasing the efficiency of management in tourism activities, in intelligent planning of tourist routes, etc.

Web 4.0 technologies are used to generate a positive experience in the minds of tourists (satisfaction) and long-term patronage (loyalty). Tourists believe that these new technologies can help improve operational efficiency, but at the same time, they want the human elements to remain; tourists like to interact with other people, they believe that cars can provide more efficient services, but human touch gives them a sense of organic happiness [34].

Considering the attitude towards the increased implementation of new technologies during pandemic, the scholars on one hand, salute the fact that technology plays a central role in the solutions for combating the COVID-19 and re-starting the tourism industry such as “mobility tracing apps, robotized-AI touchless service delivery, digital health passports and identity controls, social distancing and crowding control technologies, big data for fast and real-time decision-making, humanoid robots delivering materials, disinfecting and sterilizing public spaces, detecting or measuring body temperature, providing safety or security) while technology is seen as a panacea to our COVID-19 driven-needs to normalize surveillance, to ensure health and safety, to collect and analyze personal data for fast decision-making” [35]. On the other hand, they voiced concerns, stressing the urge for tourism research to “fight this digital trojan horse from the inside by questioning and resetting their purposes, designs and affordances, interpretations and application ethics” [35]. In this context, we were interested in exploring what is the immediate perception of the practitioners in tourism regarding the use of new technologies, especially during the pandemic.

Considering the research problem, the research objectives were set, as follows: (O1) to identify the level of familiarity of tourism operational-level practitioners from a particular destination with the concepts and use of different emergent technologies in general; (O2) to explore the level of awareness of the presence and use of the new technologies in the local practices in Brasov as a tourism destination; (O3) to identify the attitude of tourism practitioners in Brasov regarding the use of the new technologies in local tourism practices and in particular—the utility of technologies during the pandemic.

3 Methodology

Brasov is one of the most popular tourist destinations in Romania, being known, both nationally and internationally, for its views, cultural richness, best landscape for mountain tourism. The city of Braşov has many accommodation and public catering units, numerous historical and cultural objectives (The Black Church, The Council House, First Romanian School Museum, The Brasov Fortress, St. Bartholomew's Church, The Weavers' Bastion, The Mureşeni House, The White Tower, The Black Tower, The Ecaterina Gate, The Council Square, The Ethnography Museum, The Hirscher House, etc.), but also for having recreation and agreement spaces (The Tâmpa Mountain, The Solomon's Stones, The Noua Park, The Adventure Park, etc.). In the municipality of Brasov another important attraction is Poiana Braşov, a tourist resort certified as a resort of national interest, which has many hotels, restaurants, leisure centers and SPA centers, as well as mountain sports and skiing facilities (tourist routes, ski slopes and cable cars).

In the present study, we have used a mix of qualitative-quantitative approaches within the theoretical framework of Galletta and Cross [36] based on semi-structured interviews followed by a short section of close-ended items in order to capture a comparative level of familiarity of the interviewees with the emergent technologies

Table 1 Descriptive of the sample

Occupation	Number	Age (M)	Gender	Area
Front receptionist	2	25	F: 66.6% M: 43.4%	70% local (Brasov) 30% non-residents working in Brasov
Waiter	2			
Tour guide	1			
Travel agent	1			
Undergraduate students in technology and tourism local program	2			
Councilor—representative of the Brasov area of the Ministry of Tourism	1			
General manager of tourism and quality insurance training organization	1			
Guesthouse administrative manager	1			
Front office supervisor receptionist	1			

and projective techniques-based items. We aimed at including in the sample of information a wide range of operational-level practitioners accounting for Brasov tourism destination' spectrum of tourism main activities. Considering the importance of technologically savvy, well-trained future human resources in tourism and their level of familiarity with the emergent technologies in tourism, we have decided for the inclusion of two undergraduate students at the local university enrolled in the tourism program.

There were a total number of 12 specialists (Table 1) which helped us reach a theoretical saturation in results [37].

The word-association projective technique in our study was used based on Kumar et al. [38] methodological underpinning. The use of projective techniques in tourism and hospitality research was very limited until recently [39]. However, in marketing studies the word-association technique especially when investigating the image of a brand in the mental representation of clients has been very fruitful [38], and could be easily translated to tourism destination brands. We used projective techniques with the purpose of exploring also spontaneous, non-socially-desirable responses in terms of effects and representations of the subjects regarding the use of emergent technologies in tourism activities. So, in this quasi-experimental design, the word-association item required the subject to respond to each stimulus with the first word that popped-up in his/her mind. The open-ended items in the respective section had a timeframe of response of 30 s.

The final question had a role of control-item, the respondents were invited to choose and argue their preference for a modern, smart tourism destination in terms

of technology or a classical, non-automatized one in order to assess their eventual personal bias towards or against the use of emergent technologies in tourism.

4 Results and Discussion

The data obtained from the open-ended section of the interviews were analyzed following the analytic framework proposed by Galletta and Cross [36]. But, apart from basic data analysis, we imported the perspective of interpretive content analysis [40, 41] defined as a research technique extracting replicable and valid inferences from explicit data to the contexts in which has been used and extended to interview data. In the first phase, we used an open coding process in order to fragment the data and extract the principal codes and group data on thematic categories, subcategories, and properties. Later, on an iterative process, we reviewed the categories in order to refine the prior and add new categories and codes:

- (a) The ambivalent attitude in the context of Covid-19 pandemic—the use of emergent technologies in tourism—is seen simultaneously as a “coping mechanism” and “foe” for the pandemic and its aftermath from individual scale impact to local industry: *“many persons still manifest residual fear of contamination, some intelligent gadgets are welcome in reducing direct contact with infected persons in the tourism industry. On the other hand, in the pandemic period, numerous individuals in tourism industry lost their jobs, therefore these gadgets could “steal” their jobs for good”* [by automatization] (tour guide); *“the new technologies would be useful during pandemic due to the fact that allows tourist information without direct contact”* (waiter); *“[the new technologies] help tourism not to die for good”* (tourism student); *“(...) technology is everywhere and we need a break from technology”* (waiter). In some participants' opinion, the moderate stand towards the use of technologies is extended outside the pandemic context: *“My view is favorable to the use of technologies as long as it increases the efficiency of operations in tourism, but I am against replacing humans with robots.”* (travel agent).

From the travel-agencies perspective, our interviewees considered that Romania tourism industry underplayed the role of technologies as a coping mechanism of the tourism industry during the pandemic *“when other countries focused on this sector and dynamize considerably the tourism industry”* (travel agent). This perspective of under-exploitation of the potential of technologies during pandemic was back-up also by other respondents.

- (b) The technological “Safeness” or “safe-nest”. Often the perception on the implementation of the new technologies in a tourism destination as a safe-destination and comfort “certification” (not just in terms of COVID-19 risk): *“I prefer modern destinations because I know I will have a high level of safety and comfort”* (travel agent); *“I consider the role of new technologies very useful during the pandemic considering that decrease the negative interactions*

between individuals and reduces the risk of contamination with SARS-COV 2 virus” (receptionist): “I have really preferred during the period of lockdown when we were forbidden to go on holiday that the emphasis on new technologies has been more powerful” (receptionist).

When the attitude towards new technologies was the direct question, the data revealed a generally positive attitude towards the use of new technologies in the pandemic. The usual argument was that they affect consistently the development and modernization of tourism destination, insure a remote system of communication and reservation and working from home for some employees in the tourism industry, they reduce the time spent waiting for both clients and working receptionist. Contrary to the opinion that technology usually has a negative impact on social life, some informants consider the most useful technological element the WEB 4.0 „because in this way people do not lose 100% the connection between them” (receptionist).

The control item revealed that the respondents preferred going to modern destination and that they are more favorable towards the use of new technologies up to the point of considering that they were not used enough during the pandemic. One of them, and it might be considered as a diagnostic for the pandemic period during which we made „non-voluntary technological binge” in order to micromanage our professional and personal life, although he was quite familiar in daily life with technology preferred a classical destination as a holiday destination motivating that the fact that „we all need a break from technology” (waiter). This fact can represent a cognitive or emotional response to the pandemics and not a long term, although it can be a common response.

The results of the attitude assessment are consistent with the data revealed by the projective section. When people hear the syntagm „new technologies implemented in tourism industry”, they associate it more frequently with the word development followed by innovative, interactive, competition, and modernization.

The most appropriate and used technologies from the respondents’ point of view were AI and VR and a particular observation were that IoT was not among them. Other elements respondents are familiar with were virtual tours, AI guides, self-check-in/out apps, virtual assistants, smart cities, etc.

The majority of respondents associate new technologies with robotics, and a considerable amount of technologies used in tourism worldwide are not familiar and not implemented in Brasov or Romania decreasing the chance for informants to be familiar with.

On the close-ended checklist that assessed the subjects’ level of familiarity with the most common new technologies, the results revealed that 37% of the respondents were familiar with the VR technologies, 31% were familiar with AI and 16% were familiarized with both Web 4.0 and respectively IoT.

The investigation of the level of familiarity of the subjects with the presence of new technologies in tourism practice in Brasov revealed a fair level of awareness, the respondents being able to give empirical examples of their use (The first museum assistant based on artificial intelligence, Virtually Brasov app, virtual guide, VR headset, the use of online platforms, automatic facilities, safe unit BU (safe-distancing

1.5 m monitor units, air quality automatic monitor, movement-sensor for automatic doors). A significant amount of emergent technologies already in use in the destination but not yet marketed enough to acknowledge not even at the local level (holographic stand, VR haptic arc, Augmented reality app, 3D book of torture instruments and techniques recommended by Maria Theresa in *Constitutio Criminalis Theresiana* in the eighteenth century available at the History Museum in Brasov). Also, among the insufficient popularized tools was the Brasov app describing the main attractions in the destination, Virtuality Braşov, Braşov Tourism App, Braşov City App, VR Club Braşov, SAFE-UNITBV (Intelligent systems for increasing safety and comfort in the premises of the Transilvania University of Braşov), One Night Galery at festival AMURAL Braşov (Artificial intelligence in art), etc.

5 Conclusions

The findings of the study as a dominant note, regarding the positive attitude of the practitioners in tourism towards the use of new technologies in tourism, confirm those of Wirtz et al. [34], especially during pandemics, with a hint of ambivalence. The ambivalent attitude in the context of Covid-19 pandemic is due to the perspective of double-edge sword—the use of emergent technologies is perceived as a way to cope with the physically distancing and as a mental escape from reality and also as a foe, in particular as another “Trojan horse” that once some industry segments has been digitalized and employers lost their jobs during pandemics, the need for human resource in the aftermath of pandemic will be definitely more reduced. Apart from labor market perspective, at the level of personal experience, the seldom manifested ambivalence towards technology could be a temporary psychological acute reaction of response to the „non-voluntary binge” of technology that allowed us to micro-manage our professional and personal life especially during lockdown and it will not be a long-term thing, it could be quite common.

In the general positive attitude towards the use of new technologies, apart from the acknowledgement of the impact on the development and modernization of tourism destination, prevailed the mental representation that technology allowed us during pandemic to stay in touch, to have a sense of togetherness during isolation. This finding comes as a surprise, considering the common opinion that technology usually has a negative impact on social life. The general mental association between the new technologies and modernization and development of tourism destinations has been confirmed also by the projective technique, these being associated with innovation, interactivity, competition, and modernization.

In order to remain competitive and attractive for the modern tourist, modern technologies are implemented at the level of tourist destinations, of which AI and VR have been identified as the most appropriate and used so far.

The results of the study are conclusive and relevant for the stakeholders in tourism considering the development of new successful destinations. The practical implications of the results for the stakeholders imply that the abrupt and salient use of

digital services and new technologies could provoke on the short term an ambivalent response from users. From an academic point of view, this study helps strengthen existing studies on the implementation of new technologies in the tourism field and at the level of tourist destinations and provides the basis for future research. This study presents some limitations. The main limitation of this research is generated by the fact that the sampling for analysis was constituted only by practitioners in tourism. Another limitation is that the research was conducted only at the level of a single tourist destination, in one country, Romania. Future research could be extended to complete this study. The authors aim to continue research on this topic, including quantitative studies, and to investigate the presence of other variables in these studies, such as tourist preferences in Romania and other states, and expand research to other destinations.

References

1. Sigala, M.: New technologies in tourism: from multi-disciplinary to anti-disciplinary advances and trajectories. *Tourism Manag. Perspect.* **25**, 151–155 (2018)
2. Foris, D., Crihalmean, N., Panoiu T.M.: The new technologies and sustainable practices in hospitality. *Bull. Transilvania Univ. Brasov, Ser. V: Econ. Sci.* **13**(62)(2), 65–74 (2020)
3. Aldebert, B., Dang, R.J., Longhi, C.: Innovation in the tourism industry: the case of Tourism@. *Tour. Manage.* **32**(5), 1204–1213 (2011)
4. Kansakar, P., Munir, A., Shabani, N.: Technology in the hospitality industry: prospects and challenges. *IEEE Consum. Electron. Mag.* **8**(3), 60–65 (2019)
5. Law, R., Buhalis, D., Cobanoglu, C.: Progress on information and communication technologies in hospitality and tourism. *Int. J. Contemp. Hosp. Manag.* **26**(5), 727–750 (2014)
6. Law, R., Leung, R., Buhalis, D.: Information technology applications in hospitality and tourism: a review of publications from 2005 to 2007. *J. Travel Tour. Mark.* **26**(5–6), 599–623 (2009)
7. Neidhardt, J., Werthner, H.: IT and tourism: still a hot topic, but do not forget IT. *Inf. Technol. Tourism* **20**, 1–7 (2018)
8. Neuhofer, B., Buhalis, D., Ladkin, A.: A typology of technology-enhanced tourism experiences. *Int. J. Tour. Res.* **16**(4), 340–350 (2014)
9. Buhalis, D., Licata, M.C.: The future eTourism intermediaries. *Tourism Manag.* **23**(3), 207–220 (2002)
10. Pencarelli, T.: The digital revolution in the travel and tourism industry. *Inf. Technol. Tourism* **22**, 455–476 (2020)
11. Malaescu S.: Culture and civilization. Cultural tourism. Practical workbook (Cultură și civilizație. Turism cultural. Caiet de lucrări practice), Cluj Napoca, Risoprint (2018)
12. Verma, A., Shukla, V.: Analyzing the Influence of IoT in Tourism Industry. In: Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (2019). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3358168
13. Car, T, Pilepić-Stifanich, L., Šimunić, M.: Internet of Things (IoT) in tourism and hospitality: opportunities and challenges. *ToSEE Tourism South. Eastern Eur.* **5**, 163–175 (2019)
14. Natocheeva, N., Shayakhmetova, L., Bekkhozhaeva, A., Khamikhan, N., Pshembayeva, D.: Digital Technologies as a driver for the development of the tourism industry. *E3S Web Conf.* **159**(1) (2019). Retrieved from https://www.researchgate.net/publication/340124453_Digital_Technologies_as_a_Driver_for_the_Development_of_The_Tourism_Industry
15. Egger, R.: Virtual reality: an innovative sneak preview for destinations. *Virtual Reality Tourism* (2017). Retrieved from <http://www.virtual-reality-in-tourism.com/virtual-reality-an-innovative-sneak-preview-for-destinations/>

16. Coxon, M., Kelly, N., Page, S.: Individual differences in virtual reality: are spatial presence and spatial ability linked? *Virtual Reality* **20**, 203–212 (2016)
17. Nayyar, A., Mahapatra, B., Le, D., Suseendran, G.: Virtual reality (VR) & Augmented Reality (AR) technologies for tourism and hospitality industry. *Int. J. Eng. Technol.* **7**(21), 156–160 (2018)
18. Tecau, A.S., Bratucu, G., Tescasiu, B., Chitu, I.B., Constantin, C.P., Foris, D.: Responsible tourism—integrating families with disabled children in tourist destinations. *Sustainability* **11**(16), 4420 (2019)
19. Hyun, M.Y., Lee, S., Hu, C.: Mobile-mediated virtual experience in tourism: Concept, typology and applications. *J. Vacat. Mark.* **15**(2), 149–164 (2009)
20. Olsson, T., Väänänen-Vainio-Mattila, K.: Expected user experience with mobile augmented reality services. In: *Workshop of Mobile Augmented Reality, MobileHCI* (2011)
21. Malaescu, S.: Place attachment genesis: the case of heritage sites and the role of reenactment performances. In Katsoni, V., van Zyl, C. (eds.) *Culture and Tourism in a Smart, Globalized, and Sustainable World*, Springer Proceedings in Business and Economics, pp 435–449 (2020)
22. Robinson, V.M., Schänzel, H.A.: A tourism inflex: generation Z travel experiences. *J. Tourism Futures* **5**(2), 127–141 (2019)
23. Ruel, H., Njoku, E.: AI redefining the hospitality industry. *J. Tourism Futures* (2020). Retrieved from <https://doi.org/10.1108/JTF-03-2020-0032>
24. Zhang, L., Sun, Z.: The application of artificial intelligence technology in the tourism industry of Jinan. *J. Phys.* **1302**(3), 032005 (2019)
25. Gretzel, U., Sigala, M., Xiang, Z., Koo, C.: Smart tourism: foundations and development. *Electron. Mark.* **25**, 179–188 (2015)
26. Loureiro, A.: There is a fourth industrial revolution: the digital revolution. *Worldwide Hospitality Tourism Themes* **10**(6), 740–744 (2018)
27. Samala, N., Katkam, B.S., Bellamkonda, R.S., Rodriguez, R.V.: Impact of AI and robotics in the tourism sector: a critical insight. *J. Tourism Futures* (2020). Retrieved from <https://www.emerald.com/insight/content/doi/https://doi.org/10.1108/JTF-07-2019-0065/full/html>
28. Han, D., Jung, T., Gibson, A.: Dublin AR: implementing augmented reality in tourism. In: Xiang, Z., Tussyadiah, I. (eds.) *Information and Communication Technologies in Tourism*, pp. 511–523. Springer International Publishing, New York (2013)
29. Bratucu, G., Ispas, A., Chițu, I., Cismaru, L., Albu, R.G., Dincă, G., Băltescu, C., Foris, D., Demeter, T.: *Brasov—The Tourist Destination Profile*. Monograph, Transilvania University of Brasov, Brasov (2014)
30. Foris D., Florescu A., Foris T., Barabas S.: Implementation of lean techniques in destination management organizations. In: Abreu, A., Liberato, D., González, E.A., Garcia Ojeda, J.C. (eds.) *Advances in Tourism, Technology and Systems. ICOTTS 2020. Smart Innovation, Systems and Technologies*, pp. 209, 293–303, Springer, Singapore (2021)
31. Buhalis D., Amaranggana A.: Smart tourism destinations. In: Xiang, Z., Tussyadiah, I. (eds.) *Information and Communication Technologies in Tourism 2014*. Springer, Cham (2013)
32. Belli, L., Cilfone, A., Davoli, L., Ferrari, G., Adorni, P., Di Nocera, F., Dall’Olio, A., Pellegrini, C., Mordacci, M., Bertolotti, E.: IoT-enabled smart sustainable cities: challenges and approaches. *Smart Cities* **3**, 1039–1071 (2020)
33. Zhang, L., Sun, Z.: The application of artificial intelligence technology in the tourism industry of Jinan. *J. Phys. Conf. Ser.* **1302**(3), 032005 (2019)
34. Wirtz, J., Patterson, P.G., Kunz, W.H., Gruber, T., Lu, V.N., Paluch, S., Martins, A.: Brave new world: service robots in the frontline. *J. Serv. Manag.* **29**(5), 907–931 (2018)
35. Sigala, M.: Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *J. Bus. Res.* **117**(C), 312–321 (2020)
36. Galletta, A., Cross, W.: Mastering the semi-structured interview and beyond: from research design to analysis and publication. In: *Handbook of Research Methods for Tourism and Hospitality Management*. NYU Press, New York (2013)
37. Guest, G., Bunce, A., Johnson, L.: How many interviews are enough? An experiment with data saturation and variability. *Field Methods* **18**(1), 59–82 (2006)

38. Kumar, V., Aaker, D.A., Day, G.S.: *Essentials of Marketing Research*. Wiley, New York (1999)
39. Hindley, A., Font, X. (2018). The use of projective techniques to circumvent socially desirable responses or reveal the subconscious. In Nunkoo, R. (Ed.), *Handbook of Research Methods for Tourism and Hospitality Management, Handbooks of Research Methods in Management series*, pp. 202–211. NYU Press, New York
40. Krippendorff, K.: *Content Analysis: An Introduction to Its Methodology*, 3rd edn. Sage, Thousand Oaks, CA (2013)
41. Driko J., Maski, T.: *Content Analysis*. Oxford Univeristy Press (2016)

Social Media Usage Patterns and Motivations in Tourism Companies: An Outlook



Vasco Agostinho, Cristina Estevão , Paulo Duarte , and Gisela Gonçalves 

Abstract Tourism is a sector where social networks are vital for business communication. This study aims to explore the motivations and objectives of Algarve tourism businesses to use social media. A questionnaire was conducted to a sample of companies linked to the tourism sector, and 101 responses were obtained to achieve the objective. Based on the results obtained, it was concluded that most of the companies studied use social networks to communicate, especially Facebook. The main reason for this type of company to use social networks is to present services to a larger number of potential customers. The low-cost associated with social networks does not emerge in the study as a relevant factor. Companies that do not use social media do so due to a lack of interest and technical resources.

Keywords Social media · Tourism · Hospitality · Usage motivation and intensity · Digital marketing

1 Introduction

Social media result from the Web 2.0 paradigm and have transformed the mediation of information and communication. However, while there are many advantages to using

V. Agostinho

University of Beira Interior, Rua Marquês D'Ávila e Bolama, 6201-001 Covilhã, Portugal

C. Estevão (✉) · P. Duarte

NECE-UBI (Research Centre in Business Sciences), Universidade da Beira Interior, Rua Marquês D'Ávila e Bolama, 6201-001 Covilhã, Portugal

e-mail: cristina.estevao@ubi.pt

P. Duarte

e-mail: pduarte@ubi.pt

G. Gonçalves

Labcom-Communication and Arts Research Centre, Universidade da Beira Interior, Rua Marquês D'Ávila e Bolama, 6201-001 Covilhã, Portugal

e-mail: gisela@ubi.pt

social media, research at the organizational level and its impact on organizational performance have not grown as rapidly as desired [16, 25].

In the case of tourism, social media management requires advance and careful planning, ensuring that the communication targets are users of the company's networks. There are guidelines and strategies for marketers in this sector, which will help to increase the rate of user consumption [32]. Sousa et al. [31] consider that in the tourism sector, social networks offer a wide range of resources that companies can use to inform followers, promote the establishment of relationships, the promotion of interaction with customers, and the internationalization of the brand.

As tourism is an industry in continuous growth, with significant weight in developed and developing countries [10] are recognized as one of today's most relevant economic and social phenomena. Tourism is an engine of economic growth, whose impact is highly significant [9, 10, 18, 30]. It is imperative to carry out studies on the theme of social networks applied to tourism and hospitality in this context.

This research aims to provide tourism enterprises a frame of reference on the use and importance of social networks in the tourism sector in general and restoration in particular. Taking into account the role that tourism plays in the economic development of businesses [13], coupled with the lack of studies on social networks in the tourism sector in the Algarve region, this work has as a general objective to explore the practices of use of social networks by Algarve businesses linked to the tourism sector.

2 Digital Marketing and Social Networks

The digital environment has enabled bonds that allow consumers to assess whether a brand fulfills its mission and values easily. This company-customer approach is currently a ubiquitous element in consumers' lives [22].

Digital marketing, with the varied set of digital tools that it makes available to companies, facilitates commercial communication between a company and its current and potential customers.

Digital marketing has as assumptions the constant analysis of consumers/customers and their behavior to develop offers according to their needs [1]. Ryan [29] adds that thinking in digital marketing is not thinking about technologies but people. The communication between marketers and customers is at its base to generate relationships and increase sales [29].

Within digital marketing, social networks have been growing substantially affecting the business [20] assuming themselves as increasingly expressive platforms. Through them, consumers can influence other consumers by sharing opinions and experiences [21]. Thus, the analysis of social networks is currently complex since it is necessary to analyze ideas, opinions, feelings, attitudes [3].

Van Dijck [11, p. 9] argues that the various types of social media constitute a "connected communication ecosystem", interdependent, and it is impossible to define boundaries between existing social networks. However, he highlights the importance

of identifying the objectives of each of these networks to understand the different types of social groups that they can reach [11].

3 Social Networks in Tourism

With the rapid expansion of the Internet, tourism marketing has moved from passive promotion to active promotion, where much of the content is partly generated by users. The rapid changes in the tourism market and the constant changes in consumer desires continuously challenge companies in the tourism sector so that they have to be aware of the changes that occur and develop clear strategies to respond to these changes by implementing them quickly and appropriately [27].

It is known that tourists are increasingly using social media to research and plan their trips and experiences, using it also to share the experience of their final choice. Thus, marketing professionals should use social networks to obtain information about the needs and tastes of tourists and create detailed content depending on the destination or the experience to be promoted, such as the spaces or the country's culture. They should turn the information into posts, encouraging consumption and sharing experiences, either during or after them [24].

To take advantage of social networks, some companies are using them as an amplifier of word-of-mouth, a tool for direct contact with customers, a direct sales channel, and a space for social commerce [12]. In this sense, Heidemann et al. [17] determined five areas of application of social networks by companies: Research and Development (product development, market research); Marketing and Sales (marketing campaigns, word-of-mouth marketing, targeted advertising, CRM); Customer Service (customer service, after-sales support), Human Resources (recruitment, employer branding) and Internal Applications (specialized research, collaboration in virtual teams, knowledge management). The growing symbiosis between technology and tourism allows for a change in the way tourists and businesses interact. It was difficult for businesses to obtain information about tourists and their preferences for a long time without having to resort to market research. In that case, nowadays, businesses have numerous sources of information [5]. Understanding how e-tourism instruments work is central to managing marketing strategies in tourism businesses and designing policy interventions by local and national authorities [23].

The use of social media has had a marked effect in the hospitality industry [7]. An example of this effect is the increased existence of online communities where customers give their feedback regarding the market, writing comments and opinions regarding brands and their products/services [7]. Through the constant exchange of opinions about the products and lived experiences, the ease of communication and interconnection between consumers influenced new potential customers [6]. Restaurants are a paradigmatic case with customers using social media to share their experience with photos of the space and/or the meal consumed, allowing comments and suggestions from other current or potential customers [7].

Malheiro, Sousa, Ferreira [26] found that hotels' most used social media resource is posting content related to the hotel, like high-quality photos of the establishment, its surroundings, its gastronomy, and well as informative publications, related to upcoming events or even award nominations. Through the activity on their social media pages, hotels can get feedback on how their strategy is developing and whether their communication has the expected results [26].

The use of social media has expanded due to the ease of use and cost and the idea of a positive and linear correlation between its use and its business value [20].

4 Method

The population for this study includes the tourism-related companies in the district of Faro, belonging to the database provided by Informa D&B ($N = 1325$). The sample under study was selected based on the classification list of Portuguese economic activities (CAE) related to tourism (551, 552, 553, 559, 561, 562, 563, 791, 799, 9104, 9321, and 9329).

For data collection, a structured questionnaire was used based on the studies by Van Dijck [11], Glucksman [15], Kotler et al. [21], Porter [28], Ahmad [2], Al-Badi and Al-Qayoudhi [4], Bradley [8], Jacobs [19], Faustino [14] for the formulation of the assessed questions using a 5-point Likert-type scale.

The questionnaire was addressed to senior managers through direct contact, phone, or electronic mail (email). The respondents were ensured data privacy and confidentiality and were informed about the study objectives.

The questionnaire was created and made available through the Google Forms tool and sent via email to the 1250 people corresponding to the selected CAE. The questionnaires were sent out in February 2020. Due to the low response rate, two resendings were carried out during the second week of March 2020 and April 2020. After several attempts, 101 responses were achieved. Data were analyzed using the SPSS software.

5 Results

The CAE 561 (Restaurants, including mobile food service activities) was the sector with the highest percentage of responses (22.8%) followed by CAE 551 (Hotels) (20.1%), with the least represented being 553 (Camping and caravan sites) (1.0%) and 799 (Other reservation services and related activities) also with 1.0%. From the companies studied, 69.3% have no marketing department. Concerning turnover, three preferred not to answer this question. The categories with the highest number of responses were the 150 and 350 thousand euros (13.9%) and over 3 million (13.9%).

Regarding the number of employees, most companies have less than ten employees (64.4%), while only 5.0% have more than 251 employees.

Concerning the leaders' academic qualifications, it was found that most respondents have higher education training, more precisely Bachelor's degree (36.6%), followed by a Master's degree with 21.8%. As for gender, most respondents are female (53.5%). The most predominant age group is between 31 and 40 years with 30.7%, followed by 51 or more years with 19.8%. About the position held, 46.6% were managing partners, followed by the manager with 15.8%. As for seniority in the company, 44.6% have been in the company for over ten years. The category between 4 and 6 years is the second most mentioned with 21.8%, and lastly between 3 years and up to 5 years with 20.8%.

5.1 Use of Social Networks

Of the 101 companies, 92 (91.1%) use social networks. Of these, 50.0% have been using social networks for more than six years and only 9.78% between 6 months and one year. Regarding the intensity of use, 31 companies (30.7%) dedicate less than 2 h to social networks per week, while only two companies (2.0%) dedicate between 19 and 24 h. It should be noted that ten companies already dedicate more than 24 h to social networks per week.

Regarding the responsibility for the management of profiles in social networks, from the 83 responses obtained, it was possible to verify that in 22 companies, the management of social networks is done by the director (21.8%), 15 companies resort to a company employee from another department (14.9%), and lastly in 12 companies in the marketing department (11.9%). Companies favor Facebook (26.3%) and Instagram (22.4%) regarding the social networks chosen. It is noteworthy that WhatsApp is preferred by 11.5% (Table 1).

Through data analysis, it is possible to conclude that all companies with marketing departments use Facebook, and 27 use Instagram to communicate with the public. Therefore, it is observed a link between the existence of the marketing department

Table 1 Social media platform usage

Social media networks	Responses		% of cases
	N	%	
Facebook	89	26.3	96.7
Instagram	76	22.4	82.6
Twitter	10	2.9	10.9
Youtube	27	8.0	29.3
LinkedIn	23	6.8	25.0
WhatsApp	39	11.5	42.4
TripAdvisor	62	18.3	67.4
Snapchat	13	3.6	14.3
Others (8 types)	339	100.0	368.5
Total	339	100.0	368.5

and the use of social networks confirmed through the results of the nonparametric statistical test U of Mann–Whitney for a significance level of 0.05 (Table 2).

5.2 *Reasons for Joining and not Joining Social Networks*

The reasons for using social networks are diverse. According to the results in Table 3, it can be seen that the most substantial reason for joining social networks is “Introduce the company’s services to a larger number of potential customers” with an average of 4.39, followed by “Access new customers more easily” with 4.35. It is also important to highlight “Keep in touch with customers” (4.33), “Improve brand credibility and prestige” (4.18), “Communication in real-time, at any time and place” (4.13), and “Expose ideas/innovations/new services/products” (4.02).

Only nine out of 101 companies indicated no presence on social networks. The analysis of the reasons for this behavior indicates that the absence of “interest in social networks” is the strongest reason ($M = 3.44$). In contrast, the weakest reasons for non-adherence correspond to “lack of financial resources” ($M = 2.11$) and “lack of technical support to clarify doubts” ($M = 2.11$). It was verified the relationship of the three reasons for non-adherence with the highest average (lack of interest in social networks, lack of technical resources, and lack of knowledge regarding the usefulness/functionality of social networks) with the turnover of the companies not having a social media presence. We concluded that four companies that invoice less than 150 thousand euros do not adhere to social networks due to the lack of technical resources and knowledge regarding the usefulness. It was also observed that three companies with a significantly high turnover are not present in social networks because they have no interest or lack knowledge and technical resources.

The main objective of the companies that are already using social media is to promote their products/services or the brand ($N = 87$, 20.4%), followed by the touchpoint activation so that the public can interact with the company ($N = 68$; 16%), to define strategic alliances ($N = 53$; 12.4%) and to reduce research and development costs ($N = 51$; 12%). Interestingly, reducing communication costs ($N = 4$; 0.9%), achieving low-cost marketing tools ($N = 2$; 0.5%), and positioning the brand ($N = 1$; 0.2%) are the objectives by those referenced.

In fact, 85 companies claim to use Facebook to promote products/services, followed by Instagram with 73 and TripAdvisor mentioned by 59 companies. The social network where companies interact more with the public is Facebook, with 65 companies, Instagram, 57 companies, and TripAdvisor, with 47.

Table 2 Relationship between the existence of a marketing department and the use of social networks

Has marketing department	Social media network									
	Facebook	Instagram	Twitter	Youtube	Linkedin	WhatsApp	TripAdvisor			
No	59	49	3	9	10	22	37			
Yes	30	27	7	18	13	17	25			

Table 3 Reasons for joining social networks

	<i>N</i>	Mean	Standard deviation
Access new customers more easily	92	4.35	0.977
Maintaining contact with customers	92	4.33	0.939
Gather feedback from customers	92	3.97	1.104
Customer loyalty	92	3.86	1.075
Introduce the company's services to a larger number of potential customers	92	4.39	0.960
Communication in real-time, at any time and place	92	4.13	1.051
Watch the market/competitors, observe/collect information	92	3.63	1.165
Identify new business opportunities	92	3.60	1.205
Exhibit ideas/innovations/new services/products	92	4.18	1.048
Reduce research and development costs]	92	3.66	1.030
Reduce communication costs	92	4.02	1.089
Accessing funding	92	3.14	1.280
Identify new suppliers	92	3.74	1.221
Establish new partnerships with distributors	92	2.23	1.268
Define strategic alliances	92	2.86	1.314
Get low-cost marketing tools	92	2.92	1.328
Ease the process of Internationalization	92	3.09	1.255
Communicate/share views with companies in the	92	3.98	1.089
Reduce research and development costs]	92	3.87	1.169
Communicate/share views with companies in the same industry	92	2.98	1.326

6 Conclusions

Based on the results obtained, we conclude that most of the tourism-related companies studied use social networks. The social network selected by more companies in the tourism sector of this region is Facebook. With this, we can conclude that Facebook is still the social network where more companies are present despite the emergence of new options. It was expected that all companies would use the social network TripAdvisor. However, this does not happen, which may be justified because some companies are afraid of negative comments from users.

From the total of answers, 47 companies have a turnover below 150 thousand euros, and most of them have less than ten employees, so it is possible to conclude that more than half of the companies are considered SMEs. A high number of respondents have a higher education degree, which may be determinant to take advantage of the potential of social networks.

As for the profile management in these networks, it was found that most companies do not have a marketing department. The managing partner is dedicated to this function, reducing the number of companies that outsource this function. Most companies have had an online presence for more than six years, and almost all of them dedicate less than 4 h per week to social networks.

Companies state that the reasons for joining social networks are related to the presentation of the company's services to a larger number of potential customers, to be able to access new customers more efficiently, to maintain contact with them.

The most used social networks to achieve the objectives are Facebook and Instagram. However, most companies consider TripAdvisor a very important network because they can get feedback from customers through it. This result is in line with the literature review when Faustino [14] states that with the evolution of digital marketing, the marketing strategies appear to promote products and services through digital channels.

Regarding the evaluation of the performance of social networks, the most critical metrics on their social networks are the reach of publications and interaction with the public. However, they also consider essential the views on their page, the followers, and the reactions, likes, and shares.

It is also possible to conclude that the lack of interest and resources seems to be the main reason some companies do not feel attracted by social networks. Among those that use social networks, many consider that the presence on social networks is a complement to the marketing strategy, so it is crucial to have a clear strategy on social networks knowing the competitors and the final consumer. We conclude then that companies realize that it is not enough to place content on social networks but that it is essentially a strategy according to the defined objectives. This presence is a central point of the company's marketing strategy.

With these conclusions, this study contributes to the construction of scientific knowledge in the field of tourism regarding the motivations and objectives of using social networks. Thus, this study helps fill gaps in the international literature on the use of social networks by tourism businesses and contributes to an assessment of the behavior of businesses in the Algarve region, which had never been the subject of studies on the topic addressed. We also believe that the conclusions of this study may serve as a reference for entrepreneurs in the tourism sector regarding the application of social networks. Notwithstanding the contributions, this study has limitations that should be pointed out. The main one is the low response rate due to several companies being in layoff due to COVID-19, and the conclusions drawn must be interpreted with some reservations.

As the research was based on a universe of companies dedicated to the tourism sector, we suggest as future lines of research the application in other activity sectors for comparison of results.

Acknowledgements The authors would like to thank NECE—Research Unit in Business Sciences, funded by the Multiannual Funding Programme of R&D Centres of FCT—Fundação para a Ciência e a Tecnologia, under the project UIDB/04630/2020.

References

1. Adolpho, C. (2009). *Google Marketing. O guia definitivo de marketing digital*. NOVATEC; 2ª edição. Rio de Janeiro
2. Ahmad, A.: A short description of social networking websites and its uses. *Int. J. Adv. Comput. Sci. Appl.* **2**(2), 124–128 (2011)
3. Akyol, S., Alatas, B.: Sentiment classification within online social media using whale optimization algorithm and social impact theory based optimization. *Phys. A Stat. Mech. Its Appl.* **540**, 123094 (2020). Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0378437119317431>
4. Al-Badi, A.H., Al-Qayoudhi, W.S.: Adoption of social networks in business: study of users and potential users In Oman. *Int. Bus. Econ. Res. J. (IBER)* **13**(2), 401–418 (2014)
5. Amaral, F., Tiago, T., Tiago, F., Kavoura, A.: Comentários no TripAdvisor: do que falam os turistas? *Dos Algarves: Multidisciplinary e-Journal* **26**(2), 47–67 (2015)
6. Bakhshi, S., Shamma, D.A., Gilbert, E.: Faces engage us: photos with faces attract more likes and comments on Instagram. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 965–974 (2014)
7. Bilgihan, A., Peng, C., Kandampully, J.: Generation Y's dining information seeking and sharing behavior on social networking sites: an exploratory study. *Int. J. Contemp. Hospitality Manag.* **24**, 245–267 (2014)
8. Bradley, A. (2011). *Employing social media for business impact: key collective behavior patterns*. <https://www.gartner.com/doc/1538316/employing-social-media-business-impact> (December 2, 2019)
9. Brau, R., Lanza, A., Pigliaru, F. (2003). *How fast are tourism countries growing? The cross country evidence*, Working Paper CRENoS 200309, Centre for North South Economic Research, University of Cagliari and Sassari, Sardinia. Available at: <https://ideas.repec.org/p/cns/cnscwp/200309.html>
10. Chao, C., Hazari, B., Laffargue, J., Sgro, P., Yu, E.S.H.: *Tourism, jobs, capital accumulation and economy: a dynamic analysis*. Paris-Jourdan Science Economiques, Working Paper 2005-16
11. Van Dijck, J.: *The Culture of Connectivity: A Critical History of Social media*. Oxford University Press, New York (2013)
12. Dong-Hun, L.: Growing popularity of social media and business strategy. *SERI Quart.* **3**(4), 112–117 (2010)
13. Estevão, C., Garcia, A., Filipe, S.: What Are the Most Critical Factors for Competitiveness of a Tourism Destination? In: Farinha, L., Ferreira, J., Smith, H-L, Bagchi-Sen S. (eds.) *Handbook of Research on Global Competitive Advantage through Innovation and Entrepreneurship, Advances in Business Strategy and Competitive Advantage (ABSCA) Book Series*, Edited by IGI-Global, pp. 261–287 (2015)
14. Faustino, P.: *Marketing Digital na Prática*. 1ª edição. Marcador, Editorial Presença S.A., Lisboa (2019)
15. Glucksman, M.: The rise of social media influencer marketing on lifestyle branding: a case study of lucie fink. *Strategic Communication*, Elon University, 77–87 (2017). <https://www.elon.edu/u/academics/communications/journal/wpcontent/uploads/sites/153/2017/12/Fall2017Journal.pdf#page=77>
16. Hassan, S., Nadzim, S.Z.A., Shiratuddin, N.: Strategic use of social media for small business based on the AIDA model. *Procedia Soc. Behav. Sci.* **172**, 262–269 (2015)
17. Heidemann, J., Klier, M., Probst, F.: Online social networks: a survey of a global phenomenon. *Comput. Netw.* **56**(18), 3866–3878 (2012)
18. Jackson, J.: Developing regional tourism in China: the potential for activating business clusters in a socialist market economy. *Tourism Manag.* **27**, 695–706 (2006)
19. Jacobs, B.: Developing performance management in a local strategic partnership: context and issues. *Int. J. Public Sector Manag.* **22**(2), 432–444 (2009)
20. Kim, S., Koh, Y., Cha, J., Lee, S.: Effects of social media on firm value for U.S. restaurante companies. *Int. J. Hospitality Manag.* **49**, 40–46 (2015)

21. Kotler, P., Kartajaya, H., Setiawan, I.: *Marketing 3.0*. Wiley, New Jersey (2010)
22. Kotler, P., Kartajaya, H., Setiawan, I.: *Marketing 4.0*. Wiley, New Jersey (2017)
23. Labanauskaitė, D., Fiore, M., Stašys, R.: Use of E-marketing tools as communication management in the tourism industry. *Tourism Manag. Perspect.* **34**, 100652 (2020)
24. Li, C., Guo, S., Wang, C., Zhang, J.: Veni, vidi, vici: the impact of social media on virtual acculturation in tourism context. *Technol. Forecast. Soc. Change* **145**, 513–522 (2019)
25. Lovejoy, K., Saxton, G.D.: Information, community and action: how non-profit organizations use social media. *J Comput. Mediated Commun.* **17**, 337–353 (2012)
26. Malheiro, M., Sousa, B., Ferreira, L. (2019). Compreender o papel das redes sociais no comportamento do consumidor: a perspetiva do setor hoteleiro na região norte. *Revista Ibérica de Sistemas e Tecnologias de Informação* **24**(11), 299–312
27. Mistilis, N., Buhalis, D., Gretzel, U.: Future eDestination marketing. *J. Travel Res.* **53**(6), 778–790 (2014)
28. Porter, J.: *Designing for the Social Web*. Editora New Riders (2008)
29. Ryan, D.: *Understanding Digital Marketing*, 4th edn. Kogan Page, London (2017)
30. Sharpley, R. (2002). The challenges of economic diversification through tourism: the case of Abu Dhabi. *Int. J. Tourism Res.* **4**, 221–235. <https://doi.org/10.1002/jtr.378>
31. Sousa, B., Malheiro, A., Veloso, C. M.: O marketing territorial como Contributo para a segmentação turística: modelo conceptual no turismo de shopping. *Int. J. Mark. Commun. New Media. Tourism Market.* 93–116 (2019). (Special Issue 5)
32. Tan, G.W.-H., Lee, V.-H., Hew, J.-J., Ooi, K.-B., Wong, L.-W.: The interactive mobile social media advertising: an imminent approach to advertise tourism products and services? *Telematics Inform.* **35**(8), 2270–2288 (2018)

Digital Eco Tree Platform: A Proposal of an Effective Mobile Application in World Heritage Cities



Irena Travar, Zamira Acosta-Rubio, Eduardo Parra- López,
and Vidina Tais Díaz-Padilla

Abstract Continuous growth and improvement of digital solutions have created new achievements that have penetrated into all aspects of life and all industry areas. Tourism, as the industry that combines and uses products and services from a large number of other industries and sectors, has faced a lot of challenges in adopting new technological advances. In this context, tourism applications are increasingly being offered, diversifying the possibilities of destinations and improving the optimization of all their resources. Digitization of cultural heritage can greatly contribute to its preservation and promotion. In particular, World Heritage Cities (WHCs) are under the constant pressure from dynamic tourist flows, and further, the development of mobile tourism applications can help urban heritage conservation. Spain, as a country with 15 cities proclaimed WHCs, represents a valid and diverse resource for studying the development of tourism applications. Thus, this paper aims to analyze the offer and main features of tourism applications at the national level in order to identify proposals for their improvement. Finally, a Digital Eco Tree Platform for mobile applications for WHCs is proposed to increase the effectiveness of these technological solutions in tourism.

Keywords Mobile applications · World heritage cities · Sustainable tourism · Spain

I. Travar (✉) · Z. Acosta-Rubio · E. P.-. López · V. T. Díaz-Padilla
Universidad de La Laguna, 38071 San Cristóbal de La Laguna, Spain
e-mail: alu0101382388@ull.edu.es

Z. Acosta-Rubio
e-mail: zacosta@ull.edu.es

E. P.-. López
e-mail: eparra@ull.edu.es

V. T. Díaz-Padilla
e-mail: vdiazpad@ull.edu.es

1 Introduction

The concept of sustainable tourism arose in the mountain region of the Alps in the late 1970s [1]. Since then, it has become the research focus of numerous academic papers. The United Nations Educational, Scientific and Cultural Organization (UNESCO) developed the Cities Platform with the aim to encourage and facilitate the adoption of sustainable practices among cities [2]. Furthermore, UNESCO sees this as an effective way to sustainably manage cultural and natural heritage. Thus, they are currently developing a mobile application that will help “to encourage the identification, protection, preservation, and conservation of cultural and natural heritage around the world” [3].

Previous studies have also shown the importance of innovations regarding Information and Communication Technologies (ICTs) in order to promote sustainable tourism and encourage sustainable tourism practices [4–6]. Moreover, tourism and hospitality businesses largely depend on the adoption of ICTs. Thus, the use of ICTs in tourism is more a necessity than a choice.

Furthermore, it has been argued that digital solutions are one of the prerequisites for better conservation and preservation of cultural heritage [7–10]. In this context, mobile applications with innovative technological solutions can be a tool for increasing awareness about sustainable practices and, thus, developing sustainable behaviors in destinations. They can be a great source of educational materials which, in combination with audio and video materials, and augmented reality, have a possibility to stimulate continuous learning among users.

The impact of COVID-19 has greatly influenced the way tourism operates. Nowadays, tourists are more concerned about their safety, and the tourism industry needs to adapt rapidly to new circumstances. Moreover, people have become aware of the wide possibilities of technological innovations in tourism. All these give mobile applications space to receive a wide audience. This is supported by the fact that younger generations (Y and Z), who are the most technology savvy, will travel the most in the future.

There is a great variety of mobile applications in tourism. They are developed for destinations as a whole or for specific attractions. Therefore, when traveling, tourists have to install several applications to access all the necessary information, which is complicated and time-consuming. In this way, one of the most important prerequisites for wider acceptance of applications, ease of use, has not been met [11, 12].

The main aim of this study is to explore the offer of mobile applications in WHCs in Spain and, based on their content and functionality, propose a new platform that will overcome all their deficiencies, suggest ideas for their improvement, and future research areas. By doing this, we intend to help Spain continue to be an international reference in sustainable tourism, as well as an example of an e-tourism destination that manages and optimizes its on-site and digital resources in the most efficient way. Thus, the following research question arises: What is the main content that mobile

applications in WHCs for tourism purposes should incorporate in order to increase their effectiveness? From this research question, four objectives have been identified:

- Propose an effective mobile application model for WHCs.
- Investigate the offer of mobile applications in WHCs in Spain.
- Study the availability and utilities of applications in WHCs at the national level.
- Offer proposals for the improvement of the applications in WHCs.

Considering all of the above, this study helps to bridge the gap between the technological capabilities of mobile applications and their practical usefulness in WHCs as tourism destinations by creation of a unique platform that will encompass all services needed when planning, organizing, or modifying itineraries.

1.1 World Heritage Cities

World Heritage, according to UNESCO, consists of places of exceptional universal value that belong to the common heritage of humanity [13]. More specifically, the definition of the “cultural heritage” was established in the 17th UNESCO Convention, held in Paris between October and November 1972, for the Protection of the World Cultural and Natural Heritage. In this sense, it is understood that the status of “cultural heritage” can be assigned to monuments, groups of buildings, and sites [14].

The WHCs Program is one of eight programs aimed at creating a network that will facilitate cities implement Sustainable Development Goals (SDGs) [2]. Only cities that meet the strict criteria developed by UNESCO can be included in the WHCs Program [3]. In the specific case of Spain, there are 15 cities that have been declared WHCs: Alcalá de Henares (Madrid), Ávila, Baeza (Jaén), Cáceres, Córdoba, Cuenca, Ibiza, Mérida (Badajoz), Salamanca, San Cristóbal de la Laguna (Tenerife), Santiago de Compostela, Segovia, Tarragona, Toledo, Úbeda [15].

1.2 Mobile Applications in Tourism

Mobile applications evolved with the use of smartphone technology. It is notable that tourism destinations are increasingly adopting mobile applications on two different levels: attractions and destinations. In this way, the travel experience gained a new dimension. Namely, tourists are given the opportunity to visualize travel experience in advance. However, the existence of many different mobile applications and their limitations in terms of user-content interactions, the small number of official applications, and the lack of specific content [16] often discourage potential users from installing them.

Empirical research shows that there is a positive relationship between the characteristics of certain mobile applications and intention to travel. For example, attitude toward using tourism applications is a mediator between the perceived usefulness,

ease of use, and eWoM, on the one side, and the intention to visit destinations, on the other side [11]. Moreover, the usefulness of mobile applications, influenced by perceived ubiquity, informativeness, personalization, and enjoyment, is an antecedent of revisit intention [17]. In addition, perceived usefulness, ease of use, enjoyment, and price value affect consumers' behavior intention toward mobile hotel booking. Considering cultural heritage, the adoption of mobile applications technology can lead to more responsible tourism development [8], but there is a need for further integration of educational and scientific content [10].

Generally, World Heritage Sites (WHSs) in Europe, Asia, and North America invest more than other parts of the world in their promotion and marketing via mobile applications [10]. Considering Spain, the city of Salamanca, one of the 15 WHCs in the country, in cooperation with the GVAM (Accessible Virtual Guides for Museums) developed a mobile application for the city. The main innovation in this mobile application is adapted to the new post-COVID normality since audio materials for guided groups have the possibility to detect potential crowds. This application received the "Digital Tourist 2020" Tourism Offer Innovation Award [18].

1.3 Design of an Effective Mobile Application Model

The use of mobile application technology is essential in the WHSs because they can serve as a platform for educational and scientific content integration. This, further, helps to increase tourists' awareness of more sustainable behavior during their stay in a destination [10]. Moreover, cultural heritage valorization and preservation can be significantly improved through the use of digital solutions. Namely, technological innovations enable cooperation and knowledge sharing between the major stakeholders in tourism—tour operators, local, and regional authorities [8]. On the other side, the development of the tourism 4.0 influences the growth of smart tourism destinations where travelers as individuals are given a priority. They want an enriched travel experience or a customized tourism product. This can be achieved through personalized tourism applications since it has been confirmed that perceived personalization influences mobile application usefulness which, then, affects the revisit intention [17].

Concerning dominant groups of tourists, it is worth mentioning that Generations Y and Z will travel the most in the future. Also, they are the most technology savvy and the most active users of digital solutions. However, it has been found that cultural heritage lacks an appealing visual presentation which discourages these younger generations from using the applications more [8].

WHCs must be viewed as complex destinations that attract tourists not only from a cultural aspect, but the pull factor also depends on their other characteristics, like location (proximity to the sea, mountains, major emitting centers), climate, accommodation offer, etc. Thus, mobile applications should incorporate all the major tourism components with all the useful information in one "place." Several studies have proved that the usefulness of mobile applications positively influences revisit



Fig. 1 Concept of the Digital Eco Tree Platform. *Source* Developed by the authors

intentions [17], consumers’ behavior toward mobile hotel booking [12], and attitude toward using tourism applications which, further, impacts intention to visit destinations [11]. The biggest advantage of mobile applications is their accessibility regardless of location and time. In this context, they can serve to plan a trip as well as to modify itineraries.

Based on the above-mentioned existing literature, this research proposes a model of a Digital Eco Tree Platform for mobile applications for WHCs (Fig. 1). However, it could be implemented in all types of tourism destinations. In a nutshell, this concept consists of the following:

- Digital—use of online technologies for generating, storing, and processing data;
- Eco—creation of an environmentally friendly product in accordance with SDGs, more precisely SDGs 4, 5, 8, 9, 11, 12, and 17;
- Tree—a metaphorical presentation where the tree is viewed as a tourism system of a city, and tree branches as a representation of major tourism components;
- Platform—more specifically platform as a service (PaaS) that implies the creation of an application available in application stores.

Furthermore, the proposed mobile application model encompasses three main actors: users, destinations, and firms. Users are given the opportunity to have all the necessary information for planning and realization of the trip. Moreover, they can be supported by real-time information. Destinations can use big data technology to gather and analyze tourists’ preferences and minimize negative environmental effects. Finally, firms can offer their products and services and generate value. The

sustainable component permeates all three operational levels, which is essential for tourism development in WHCs [19].

The main goal of this research was to propose an application that can satisfy tourists' essential needs when traveling, and these are transportation, accommodation, and food. Besides that, the application should offer additional services in the destination, like attractions, recreational activities, events and festivals, and shopping. In this way, tourists will have everything they need, when planning, organizing and, eventually, modifying an itinerary in one "place."

2 Methodology

Regarding the methodology used, it consisted of three main steps. After developing a brief theoretical framework to clarify the basic concepts and the first approach to an effective mobile application model—the Digital Eco Tree Platform, data collection was conducted. Google Play Store, that supports the Android mobile operating system, was used to find them mobile applications in WHCs in Spain. The keywords "[*The name of the city*] tourism" were entered into the search engine, and only applications labeled "Travel & Local" were included in the analysis. Another prerequisite was that applications were developed for cities as destinations. Moreover, applications in both English and Spanish were considered. At least, one application was found for each city.

After this first step, a comprehensive search of applications' content and functionality [10] was performed as well as the extant literature review. For each application, the main characteristics in terms of content and functionality were identified, which gave us a unique basis for establishing indicators. This determined a set of seventeen indicators [10, 16, 20] that covered the main applications' features, multimedia, and tourism information. If an indicator describing a particular feature was presented in the application, the value 1 was placed in the grid, while the value 0 indicated its absence.

The third step consisted of comparing applications according to established indicators and identifying gaps between the possibilities of mobile technology and their exploitation in relation to mobile applications in WHCs in Spain. This is of great importance for the configuration of the efficient mobile application proposal launched in this paper, based on a comprehensive literature review and targeted at various public or private national institutions.

3 Results

The research revealed that, for Android, there was a total of 48 mobile applications for 15 WHCs in Spain. All applications were downloaded and analyzed. However, it was found that three applications had certain issues and information could not be

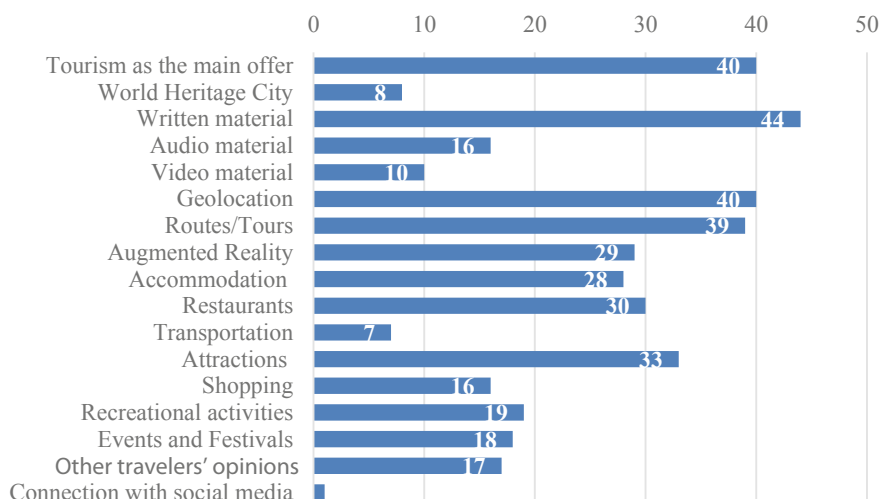


Fig. 2 Presence of indicators for Android mobile applications for World Heritage Cities in Spain.
Source Own elaboration

retrieved. Thus, 45 applications were included in the final analysis. It is important to highlight that all applications were free of charge. The analysis included a comparison of the main applications' features, multimedia, and tourism information. Besides that, applications were classified in relation to the number of installs, rating, and the number of languages in which each application was available.

Figure 2 illustrates the presence of indicators for Android mobile applications for WHCs in Spain. Even though only the applications labeled "Travel & Local" were included in the research, it was found that tourism was the main offer in 40 applications (88.9%). Also, it was noted that all the analyzed applications had "Written content" which referred to explanations of certain routes and attractions. Other most frequently present indicators regarding application features and multimedia were: "Geolocation" (88.9%), "Routes/tours" (86.7%), and "Attractions" (73.3%). In addition, the indicators that were present in more than half of the analyzed applications were: "Restaurants" (66.7%), "Augmented Reality" (64.4%), and "Accommodation" (62.2%). The indicator that was the least present was "Connection with social media." Only one application had this feature. The indicator related to "Transportation" was barely present (15.6%) as well as the details regarding WHCs, like UNESCO logo or any written information (17.8%).

The number of installs for Android mobile applications for WHCs in Spain is presented in Fig. 3. It can be noted that 12 applications (26.7%) have less than 1000 downloads, of which as many as four (8.9%) have less than 50 installations. The highest number of applications ($n = 22$, 48.9%) have between 1000 and 5000 downloads. However, more than 5000 users downloaded six (13.3%) out of 45 analyzed applications, while only five applications (11.1%) have more than 10,000 installations.

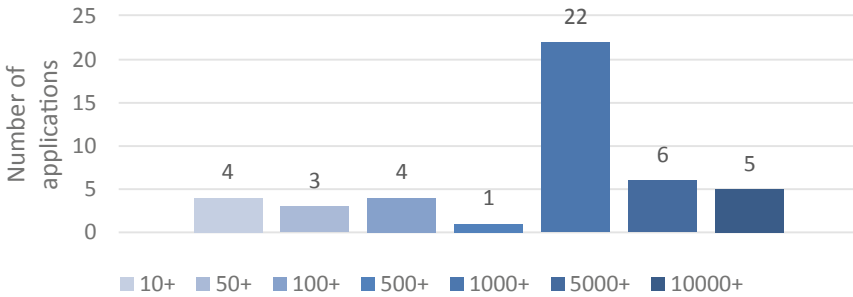


Fig. 3 Number of installs for Android mobile applications for World Heritage Cities in Spain. *Source* Own elaboration

Regarding the rating of the applications, it went from 2.9 to 5.0. The score below 3.0 had only three applications (6.7%). The rating below 3.1 and 4.0 had 10 applications (22.2%), while the rating higher than 4.1 had 20 applications (44.4%). It is worth mentioning that as many as 12 applications (26.7%) had no reviews.

Considering the number of languages, all analyzed applications offered content in Spanish, of which two-thirds had written, audio, or visual materials exclusively in this language.

4 Discussion

The current study aimed to analyze mobile applications for WHCs in Spain as well as to take a deeper understanding of potential drawbacks that impede their wider adoption. In this regard, an effective model for mobile applications was proposed.

Some of our findings are consistent with previous studies. Namely, it was found that geolocation and information about tours and attractions are widely presented in mobile applications. This is in line with the research done in several tourism destinations in Spain, according to which geolocalization and the city resource guide are the most prominent functions, while sharing comments and the valuation of places do not frequently appear [16]. However, we concluded that the information about events and festivals was presented in only half of the analyzed applications, and a large number of applications used augmented reality. This is not consistent with the previous study in Spain that demonstrated the opposite findings [16]. In addition, emphasizing the protection of the city as a WHC, in the form of a logo, written, audio, or video information, did not find a wide presence among the analyzed applications. Similarly, previous research on UNESCO WHSs revealed that “the WHS logo and the UNESCO logo were present only in 12.2% and 11.3% of applications, respectively” [10, p. 24].

Past research has indicated that tourists are more motivated by consumption experience than by activities related to sightseeing and leisure. Thus, the suggested type of

content for mobile applications is information about restaurants, pubs, and shopping [20]. Our findings do not support these recommendations since only 16 applications (35.6%) provided information about shopping, while tour and attractions suggestions were the ones of the most frequently found indicators.

The lack of the wide adoption of mobile applications could be explained by the fact that young people, who are the most active technology users, do not find the heritage content appealing [8]. The content analyzed for the purposes of this research is very similar to the one on the official websites, and people prefer a unique and specific content [16]. Also, one of the possible reasons could be that mobile applications do not provide comprehensive information. Instead, potential users need to install different applications if they want, for example, the information about local tours, gastronomy, accommodation, transport, shopping, etc. In addition, only 15 mobile applications (33.3%) offered content in two or more languages, even though Spain was the second country with the largest number of international tourist arrivals worldwide in 2019 [21].

Considering the above-mentioned discussion, the model application proposed in this research would overcome all the barriers of the existing mobile applications in terms of tourists' needs. Namely, tourists would have all the information they need in one mobile application when planning, organizing and, eventually, modifying an itinerary. And this is something that all the analyzed applications lack. Furthermore, it would integrate all the important characteristics for their wider adoption, like ubiquity, informativeness, personalization [17], enjoyment [12, 17], usefulness, ease of use [11, 12], and eWoM communication [11]. Certainly, the creation of the Digital Eco Tree Platform is a complex process that requires the involvement of all tourism stakeholders, but this is a place where new digital solutions, like big data and metadata, can come to the fore.

5 Conclusion

Digitalization of cultural heritage is an essential aspect of its future promotion and preservation. In this context, it is important to find a linkage between new tourism trends and technological achievements. Regarding mobile applications for tourism purposes and their still modest adoption, the following question arises: "What do tourists, and especially younger generations, expect mobile applications should have to satisfy their travel needs and in-situ experiences?"

This study explored the academic literature concerning the adoption of mobile applications to find the answer to this question. Moreover, the current offer of mobile applications in WHCs in Spain, with an emphasis on their content and functionality, was evaluated. Hence, the model of an effective mobile application was developed that would encompass digitalization, sustainability, major tourism components, and a PaaS. It is based on ease of use and usefulness, and the creation of this platform

would allow the identical model to be used in all WHCs worldwide. In addition, it would overcome all the barriers that impede a wider adoption of mobile applications in tourism.

6 Limitations and Future Research Directions

Regarding limitations, this study disregards mobile applications developed for other operating systems different from Android, which may have led to the inclusion of a limited number of applications in the research. Moreover, the lack of empirical research significantly affects the outcome of the proposal for an effective mobile application. Namely, users' preferences in specific locations, like WHCs in Spain, were not assessed. This caused that only the main content and general features of the proposed effective mobile application could have been suggested.

In the future, more profound research should be undertaken that will take into consideration mobile applications for other operating systems, like iPhone. Furthermore, empirical research would help to further develop the concept of the Digital Eco Tree Platform and to thoroughly define the content, functionality, utility, and usability of the proposed model. Since this study refers to WHCs in Spain, it would be beneficial to conduct the same theoretical and empirical research in other countries that would allow the comparison of findings. Finally, the widespread adoption of mobile applications for tourism purposes can have a positive impact on the achievement of various SDGs. Hence, further research toward this direction is suggested.

References

1. Bramwell, B., Lane, B.: Sustainable tourism: an evolving global approach. *J. Sustain. Tour.* **1**(1), 1–5 (1993). <https://doi.org/10.1080/09669589309450696>
2. “UNESCO for Sustainable Cities,” unesco.org. <https://en.unesco.org/unesco-for-sustainable-cities>. Accessed 21 Sept 2021
3. “Towards smart cities,” unesco.org. <https://en.unesco.org/courier/2019-2/towards-smart-cities>. Accessed 21 Sept 2021
4. Ali, A., Frew, A.J.: ICT and sustainable tourism development: an innovative perspective. *J. Hosp. Tour. Technol.* **5**(1), 2–16 (2014). <https://doi.org/10.1108/JHTT-12-2012-0034>
5. Gössling, S.: Technology, ICT and tourism: from big data to the big picture. *J. Sustain. Tour.* **29**(5), 849–858 (2020). <https://doi.org/10.1080/09669582.2020.1865387>
6. Guemide, B., Benachaiba, C., Maouche, S.: Integrating ICT-based applications for sustainable tourism development in Algeria. *J. Tour. Hosp.* **8**(5), 1–15 (2019). <https://doi.org/10.35248/2167-0269.19.8.415>
7. Cantoni, L.: Heritage and sustainable tourism. The role and challenge of information and communication technologies. In: Silvia De Ascaniis, L.C., Gravari-Barbas, M. (eds.) *Tourism Management at UNESCO World Heritage Sites*. Università della Svizzera italiana, Lugano, pp. 67–73 (2018)

8. Varbova, V., Zhechkov, R.: Digital solutions in the field of cultural heritage (2018). [Online]. Available <https://ec.europa.eu/culture/sites/culture/files/library/documents/staff-working-document-european-agenda-culture-2018.pdf>
9. Permatasari, P.A., Qohar, A.A., Rachman, A.F.: From web 1.0 to web 4.0: The digital heritage platforms for unesco's heritage properties in Indonesia. *Virtual Archaeol. Rev.* **11**(23), 75–93 (2020). <https://doi.org/10.4995/var.2020.13121>
10. Schieder, T.K., Adukaite, A., Cantoni, L.: Information and communication technologies in tourism 2014. In: *Information and Communication Technologies in Tourism 2014* (2013, January). <https://doi.org/10.1007/978-3-319-03973-2>
11. Kuo, T.S., Huang, K.C., Nguyen, T.Q., Nguyen, P.H.: Adoption of mobile applications for identifying tourism destinations by travellers: an integrative approach. *J. Bus. Econ. Manag.* **20**(5), 860–877 (2019). <https://doi.org/10.3846/jbem.2019.10448>
12. Mohamad, M.A., Radzi, S.M., Hanafiah, M.H.: Understanding tourist mobile hotel booking behaviour: incorporating perceived enjoyment and perceived price value in the modified Technology Acceptance Model. *Tour. Manag. Stud.* **17**(1), 19–30 (2021). <https://doi.org/10.18089/TMS.2021.170102>
13. “Patrimonio Mundial,” unesco.org. <https://es.unesco.org/themes/patrimonio-mundial>. Accessed 22 Sept 2021
14. UNESCO, Convención sobre la protección del patrimonio mundial, cultural y natural (1972)
15. whc.unesco.org, “World Heritage List.” <http://whc.unesco.org/en/list/&order=country>. Accessed 22 Sept 2021
16. Fernández-Cavia, J., López, M.: Communication & society. *Commun. Soc.* **26**(2), 95–113 (2013)
17. Briliana, V., Prasetyo, A.B.: Some antecedents and effects of using mobile apps in tourism marketing. In: *Proceedings of the 7th International Conference on Entrepreneurship and Business Management (ICEBM Untar 2018)*, pp. 270–275 (2019, January). <https://doi.org/10.5220/0008491702700275>.
18. gvam.es, “La digitalización del turismo hacia las ciudades inteligentes,” (2021). <https://www.gvam.es/digitalizacion-turismo-ciudades-inteligentes/>. Accessed 19 Sept 2021
19. Kim, D., Kim, S.: The role of mobile technology in tourism: patents, articles, news, and mobile tour app reviews. *Sustainability* **9**, 2082 (2017). <https://doi.org/10.3390/su9112082>
20. Rivera, M., Croes, R., Zhong, Y.S.: Developing mobile tourism services for a destination: the case of Aruba. vol. 28, no. 12, pp. 2721–2747 (2016). <https://doi.org/10.1108/IJCHM-02-2015-0052>
21. “Countries with the largest number of international tourist arrivals worldwide in 2019,” statista.com, 2021. <https://www.statista.com/statistics/261726/countries-ranked-by-number-of-international-tourist-arrivals/>. Accessed 21 Sept 2021

Virtual and Augmented Reality in Marketing

An Integrated Framework of the Relation Between Augmented Reality and Brand Love



Carolina Afonso and Vanessa Hipólito

Abstract Augmented reality (AR) technology is becoming increasingly important not only to consumers, but also to brands and businesses. The growing demand for unique and innovative experiences is facilitating its fast growth. One of the current market challenges is to understand how AR technology can develop consumer–brand relationships. Thus, this paper aims to propose a conceptual model that helps to better understand the relation between the usage of augmented reality and the development of brand love. Qualitative methodology was utilized in this study. Data were collected through a focus group, and content analysis was performed. Results showed that there is a positive correlation between the usage of AR and the development of brand love. At the academic level, this paper presents an integrated framework about the relation between augmented reality, brand love, and its dimensions. At the enterprise level, it provides key insights to managers and marketers that enable an alignment of the brand–consumer relationship through the incorporation of AR.

Keywords Augmented reality · Brand love · Marketing · Content analysis

1 Introduction

The growing consumer demand for experiences enriched with virtual content and the development of marketing strategies integrated with augmented reality (AR) technology is beginning to gain prominence [1]. AR allows for opportunities to interact with consumers and personalized experiences [1, 2]. AR should be recognized as an essential tool, not only for the consumer but also for marketing in general.

Technological evolution brings consumer access to a wide range of brands and products [3]. This creates concern on the part of brands and companies to attract and

C. Afonso (✉)

Advance/CSG, ISEG, Universidade de Lisboa, Lisboa, Portugal

e-mail: carolinaafonso@iseg.ulisboa.pt

V. Hipólito

ISEG, Universidade de Lisboa, Lisboa, Portugal

retain their consumers. When seeking to establish closer and more emotional connections with the consumers [4], a concern regarding brand management concepts, such as brand love, arises [5–8]. When consumers are required to decide between similar brands, they increasingly start to use criteria such as love [9]. It has been shown that brands commanding high levels of love have managed to increase their value by 191% over a 12-year period. Therefore, it is considered a contemporary challenge for brands and companies to understand the association between the use of AR technology and the development of the relationship between the consumer and the brand. The general objective of this paper is to identify and understand these aspects.

2 Literature Review

2.1 *Augmented Reality (AR): Definition*

AR technology emerged in cinematography in 1950 by Morton Heilig [10]. Recent technological developments, combined with the growing use of smartphones and other mobile devices, contributed to the growth of AR importance worldwide. AR technology is characterized by the ability to modify and improve the physical environment by transmitting information or virtual objects generated by a computer [6], interactively and in real time [10]. Consistently defined by the most diverse authors [10–12], it differs from other interactive technologies as it enables real-time interaction [2]. Despite being considered an aspect of virtual reality (VR), AR differs from VR as its immersion in the virtual world is not complete [1]. AR allows the user to simultaneously observe the real world [13] while interacting with the virtual elements [2, 10].

Widely investigated over the last few years (especially in the marketing area), AR is beginning to gain importance in different industries and sectors. More and more brands and companies are integrating AR technology into their current marketing strategies [1]. Although there is still no definition of AR marketing [1], it is designated as a strategic concept where physical and digital information is combined, benefiting the consumer, brands, and companies. With AR, real and virtual objects and information can be superimposed and aligned with each other [14]. This is possible through the usage of mobile devices such as smartphones, tablets, wearables, interactive screens or projectors, and more recently, through AR applications [15]. Within the wide range of AR devices and applications, smartphone front cameras have developed the AR mirror feature, which embodies the most advanced version of virtual, real-time try-on [16]. The AR mirror feature allows the user to try or test products through the image on the screen with superimposed virtual elements [16].

AR allows marketers to develop immersive narratives, facilitating a new product experimentation experience for consumers as well as developing personalization, leading to an increase in satisfaction [17], assisting in decision making, and encouraging purchase [18].

2.2 AR Characteristics

According to several authors, AR technology is composed of three key characteristics that make it distinctive from other technologies: (1) the ability to combine real and virtual objects; (2) real-time interactivity; and (3) the ability to record in 3D with regard to the alignment between objects, virtual information, and information from the real environment [2, 10, 11, 13].

Herein, three key AR characteristics are adopted: (1) augmentation, (2) interactivity, and (3) registration [13].

Augmentation. AR technology allows for the increase or overlap of the physical environment with virtual elements. This is considered the main differentiator of this technology [11]. This characteristic, called magnification [13], is defined as the superposition of virtual objects in real environments with the objective of adding or hiding part of the real environment. It can occur through different virtual elements such as text, images, videos, audio, or geolocation information [19]. Authors [13] add that this characteristic not only refers to the sense of sight, but also to hearing, taste, touch, and smell. The augmentation can occur over a person, product, or surrounding space [11]. Among the various AR devices and applications, AR mirrors stand out, allowing the user's image to be modified by generating enlarged images [11]. Consumers perceive the augmentation related to their image and consequently relate to the products, as well as the resulting behavioral intentions [11].

Interactivity. Several authors have identified interactivity as one of the main characteristics of AR technology [13, 18]. Although the augmentation is considered the distinguishing feature of AR technology, the success of its performance requires that it be interconnected with the resource of interactivity [1]. It is through this characteristic that interactive technologies produce 3D content, with different directions, styles, and colors, where the degree of interaction actively manipulates the provided virtual information [1]. Widely studied in recent times, there are several definitions of the concept of interactivity. Authors argue that there are two different approaches to interactivity: (1) based on the characteristics of the device and (2) based on the user's perspective [20]. From the user's perspective, interactivity allows them to develop immersive experiences and generate affective, cognitive, and behavioral responses. Interactivity allows consumers to experience the products as if they were trying them in-person [17, 20].

Registration. In order to ensure that the connection between the physical and real worlds will not be compromised, the objects and virtual information superimposed on the real environment must be perfectly aligned [13]. Information and virtual objects must have a specific location and should be able to be surveyed from any point of view, changing the perspective of the respective virtual increase with the user's movements and must have a minimal processing time [10].

2.3 AR and the “Consumer–Brand” Relation

Some authors identified the ability of AR technology to develop positive relationships between consumers and the brand [17]. Through the usage of AR applications, they can develop more intimate relationships with their consumers, rather than merely transactional. According to the same authors, this technology can be analyzed in two different contexts: immediate and internal. The immediate context puts the alignment between virtual and real objects in the foreground, whereas the main concern is the analysis of the consumers’ immediate responses to applications based on transactional relationships [17]. They also identified that the internal context becomes significantly more relevant in the analysis of relationships, being directly related to the potential impact on the consumers’ sense of identity. They recognized the ability of AR technologies to offer authentic experiences to the consumers, and consequently, a superior added value, creating and preserving an emotional connection between the consumer and the product or brand [2].

Using self-improvement applications such as AR mirrors, it is possible to augment a person [11]. The theory developed by Belk [21] about the “augmented self” supports these findings. Users employ their consumption experiences to build their own identity, consequently triggering positive emotions [21]. In this way, the “increase of the self” triggers a set of positive feelings not only for the brand itself, but also for the user. AR’s ability to develop brand love based on the theory of self-reference by identifying the AR environment is more suitable for its development in the face of environments that do not have it [22]. Brand love is developed based on a sense of identity shaped by simulation recorded by technology, as well as a sense of control and execution. Other authors [16] used AR mirrors to identify that the holistic (rather than partial) increase affects the way consumers perceive themselves, and consequently, their self-esteem. They also concluded that this holistic augmentation affects the way consumers understand the product with regard to its impact on their self-perception. Specifically, the use of AR mirror allows the establishment of a relationship with these products resulting in a feeling of an “augmented me” [16]. Additionally, the sense of self-identity, developed through the application of AR, is the main driver in the development of deeper relationships. Although the association between AR and brand love has been researched, the existing literature on this association is still very limited [22]. There are no known studies that have specifically analyzed the association between the use of AR applications and the development of brand love, although several authors have identified the capacity of branding to develop closer relationships between these aspects [2, 18]. Thus, it is considered relevant to analyze the association between the use of AR applications and the development of brand love. The following research question must be analyzed:

RQ1. Does the use of AR applications allow for the development of brand love?

2.4 *Brand Love Definition*

The concept of brand love is currently gaining importance among academics and marketers [7]. Despite the extensive studies carried out over the past few years, [5–7] there is still no consensus on its definition [7]. Some brands manage to inspire feelings and emotions in their consumers through objects and consumption activities [5]. However, regardless of the emotional attachment and feeling of love shown by some consumers toward objects or activities, this is not fully analogous to interpersonal human love [5]. Brand love is the degree of emotional attachment, passion, positive evaluation, positive emotions, and declarations of love that a satisfied consumer nurtures for a brand after the consumption experience [6]. They feel loyalty toward the brand and share a favorable opinion of it through word-of-mouth. Additionally, they also verify that the hedonic product category and the ability to integrate the brand into the consumer's identity have a satisfactory effect on the development of brand love [7]. Contrary to these studies, a different concept of brand love has been developed. Based on consumer behavior and the way they understand and experience the concept (instead of assuming interpersonal human love as a basis), they define brand love as an adaptation of the consumer to the consumption situation [7]. They also define the relationships between consumers and brands as non-altruistic on the part of the consumer, and not reciprocal on the part of the brand.

2.5 *Brand Love Dimensions*

A framework for measuring brand love has been proposed, which identifies six distinct core elements of brand love [7]. First, private brand integration is the central aspect of brand love and is closely related to the construction of the consumer's current or desired identity through a direct relationship with the object [23]. Second, the behaviors driven by passion are reflected by the strong desire for the brand and their products [24] with a willingness to invest, regardless of the price [25]. Third, a positive emotional connection is established between the consumer and the brand, where consumers consider it unique and irreplaceable [26]. Fourth, separation anxiety is experienced as the consumer gains a strong desire to maintain a relationship with the brand [7], suffering in advance for its eventual separation or disappearance [7]. Fifth, a permanent relationship is generated through the long-term commitment between the consumer and the brand. Sixth, positive valence is formed in relation to the satisfaction of the consumer [7]. Thus, the following proposition is formulated and the conceptual model is proposed (Fig. 1).

P1. There is a positive and significant relation between AR and the development of brand love.

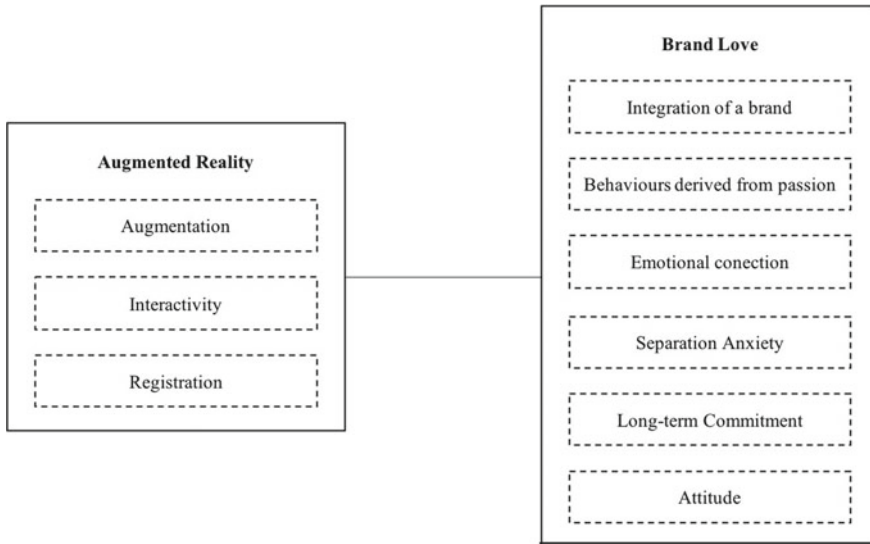


Fig. 1 Conceptual model

3 Methodology

The analysis of the present study was carried out through a qualitative approach. Data were collected through a focus group lasting 1 h and 30 min. The sample consisted of a total of six female participants aged 19 through 60 years, all of whom resided in Portugal and had previously used AR applications.

4 Results

Content analysis was performed, and eight categories were created regarding the characteristics of AR most valued by the participants. Among these categories, AR dimensions (characteristics related to registration and interactivity) were identified as the most significant in the development of an AR application (Fig. 2). For the participants, it is essential to ensure that the application conveys confidence and contains an experience that is close to reality through a good level of registration. However, it is equally important to be able to, through the application's interactive capability, conjugate, and interact with products in real time.

Through the analysis of the participants *verbatim*, the use of augmented reality applications allowing for the development of brand love was discovered (Table 1). A positive association was identified between the use of AR applications and: (1) the integration of a brand, both at the level of current self-identity and at the level of desired self-identity [11, 16]; (2) behaviors derived from passion, as it awakens the

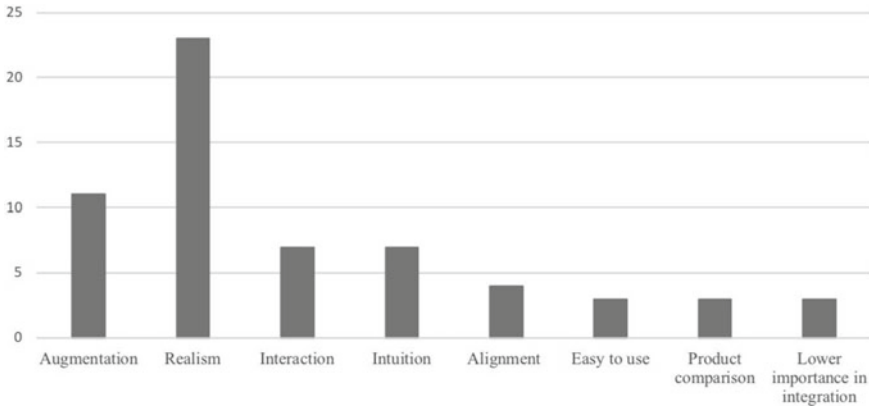


Fig. 2 Frequency of the categories related to AR characteristics

Table 1 *Verbatim* regarding Brand Love

Sociodemographic characterization	<i>Verbatim</i>
22 years, Female, Degree	“The application allowed me to see, at the end, an image of the product that I identified myself a lot with”
55 years, Female, Degree	“Through the AR feature of the application, it was possible to combine different colors of lipsticks with different colors of eye shadows that helped to make my choice. I feel now more close to the brand”
23 years, Female, Degree	“The applications with AR allow you to try the products and evaluate them better. It is a key benefit against a brand that does not have this technology”

desire to buy the products experienced; (3) emotional connection, with the ability to develop a superior connection with the brand and its products [2]; (4) separation anxiety, where the possible disappearance of an AR application arouses a negative impact; and finally, (5) the attitude, increasing empathy for the brand and its products [2]. However, there is no evidence that supports the theory that the dimension of long-term commitment has a positive association with the use of AR applications.

5 Conclusions, Limitations, and Further Research

This paper was developed with the objective of proposing a framework in order to explore and deepen the knowledge about AR technology and applications, and the development of brand love. Specifically, its objective was to understand and analyze the association between the use of these applications and the possible development

of relationships between the consumer and the brand. There were multiple significant characteristics that constitute AR applications. This study also allowed us to understand the association between the use of AR applications and the development of relationships between the consumer and the brand, namely the development of brand love.

At the academic level, the paper thus allows the development of a conceptual model, establishing new relationships that had not been previously studied, thus filling a gap in the existing academic literature. At a business level, the analysis of the various topics explored in this research becomes particularly relevant due to the current emergence and dimension that this technology is beginning to gain, not only among consumers, but also among brands and companies. Through the analysis of consumers' perception of AR technology, it is possible to conclude that the main benefits taken into account in an AR application are the possibility of product experimentation, as well as the opportunity to make the brand and its products more widely known. Therefore, the development of experiences using AR, focusing on these main topics, should be a major area of research. Companies can now start to focus on these three characteristics: augmentation, interactivity, and registration.

The present study had certain limitations: the use of non-probabilistic samples and a sample composed only by female participants. Thus, in future research, a study with a quantitative approach, with a broader and diversified sample, should be developed. Finally, other types of relationships between the consumer and the brand, such as brand activation or awareness, should also be explored.

Acknowledgements We gratefully acknowledge financial support from FCT—Fundação para a Ciencia e Tecnologia (Portugal), national funding through research grant UIDB/04521/2020.

References

1. Rauschnabel, P.A., Félix, R., Hinsch, C.: Augmented reality marketing: how mobile AR-apps can improve brands through inspiration. *J. Retail. Consum. Serv.* **49**, 43–53 (2019)
2. Kiryakova, G., Angelova, N., Yordanova, L.: The potential of augmented reality to change the business. *Traika J. Sci.* **15**(1), 394–401 (2017)
3. DeMers, J.: 7 Technology trends that will dominate 2018. <https://www.forbes.com/sites/jaysondemers/2017/12/30/7-technology-trends-that-will-dominate-2018/#19b49e0357d7>. Last accessed 2 Sept 2021
4. Lima, M. J., Pinto, M.J.: Fórum Comunicação: Proximidade gera relações duradouras. *Marketeer* 100–104 (2018)
5. Ahuvia, A.C.: Beyond the extended self: loved objects and consumers' identity narratives. *J. Consum. Res.* **32**(1), 171–184 (2005)
6. Carroll, B.A., Ahuvia, A.C.: Some antecedents and outcomes of brand love. *Mark. Lett.* **17**(2), 79–89 (2006)
7. Batra, R., Ahuvia, A., Bagozzi, R.P.: Brand love. *J. Mark.* **76**(2), 1–16 (2012)
8. Bagozzi, R.P., Batra, R., Ahuvia, A.: Brand love: development and validation of a practical scale. *Mark. Lett.* **28**(1), 1–14 (2017)
9. Caras, A.: What builds brand love. In: Paper presented at the Annual Meeting of the IAB Europe Interact 2018, Milan, Italy (2018)

10. Craig, A.B.: *Understanding augmented reality: concepts and applications*, 1st edn. Waltham, M. A., Elsevier (2013)
11. Javornik, A.: Augmented reality: Research agenda for studying the impact of its media characteristics on consumer behaviour. *J. Retail. Consum. Serv.* **30**, 252–261 (2016)
12. Poushneh, A., Vasquez-Parraga, A.Z.: Emotional bonds with technology: the impact of customer readiness on upgrade intention, brand loyalty, and affective commitment through mediation impact of customer value. *J. Theor. Appl. Electron. Commer. Res.* **14**(2), 90–105 (2019)
13. Azuma, R.T.: A survey of augmented reality. *Teleoperators Virtual Environ.* **6**(4), 355–385 (1997)
14. Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. *IEEE Comput. Graph. Appl.* **21**(6), 34–47 (2001)
15. Javornik, A.: It's an illusion, but it looks real! Consumer affective, cognitive and behavioral responses to augmented reality applications. *J. Mark. Manag.* **32**(9–10), 987–1011 (2016)
16. Javornik, A., Pizzetti, M.: "Mirror, mirror on the wall, who is real of them all?" The role of augmented self, expertise and personalization in the expertise with augmented reality mirror. *Adv. Consum. Res.* **45**, 423–427 (2017)
17. Poushneh, A.: Augmented reality in retail: a trade-off between user's control of access to personal information and augmentation quality. *J. Retail. Consum. Serv.* **41**, 169–176 (2018)
18. Poushneh, A., Vasquez-Parraga, A.Z.: Discernible impact of augmented reality on retail customer's experience, satisfaction and willingness to buy. *J. Retail. Consum. Serv.* **34**, 229–234 (2017)
19. Pescher, C., Reichhart, P., Spann, M.: Consumer decision-making processes in mobile viral marketing campaigns. *J. Interact. Mark.* **28**, 43–54 (2014)
20. Oh, J., Sundar, S.S.: How does interactivity persuade? An experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *J. Commun.* **65**(2), 213–236 (2015)
21. Belk, R.W.: Possessions and the extended self. *J. Consum. Res.* **15**(2), 139–168 (1998)
22. Huang, T.-L.: Psychological mechanisms of brand love and information technology identity in virtual retail environments. *J. Retail. Consum. Serv.* **47**, 251–264 (2019)
23. Holt, D.B.: Poststructuralist lifestyle analysis: conceptualizing the social patterning of consumption in postmodernity. *J. Consum. Res.* **23**(2), 326–350 (1997)
24. Belk, R.W., Ger, G., Askegaard, S.: The fire of desire: a multisited inquiry into consumer passion. *J. Consum. Res.* **30**(3), 326–351 (2003)
25. Thomson, M., MacInnis, D.J., Park, C.W.: The ties that bind: measuring the strength of consumers' emotional attachments to brands. *J. Consum. Psychol.* **15**(1), 77–91 (2005)
26. Fournier, S.: Consumers and their brands: developing relationship theory in consumer research. *J. Consum. Res.* **24**(4), 343–373 (1998)

A View of Augmented Reality in the Beauty Industry from an Exploratory Perspective: Generations X and Z



Sergio Barta, Raquel Gurrea, and Carlos Flavián

Abstract Augmented reality (AR) is shaping a new environment for marketers, playing an essential role in consumer choice. AR seems to be a good solution for undecided online shoppers. However, their implementation in the physical store can also play an important role in both the purchase decision and in increasing the number of consumers visiting the shop. AR allows potential consumers to virtually try-on products and imagines how they would look on their face, through technologies such as Virtual Try-on. From a qualitative perspective, this research aims to explore perceptions of AR and then differentiate between the use of Virtual Try-on from the smartphone and the use of Virtual Try-on in the physical shop through mirrors. Through two focus groups, composed of women from two generations (generation X and generation Z), we found a greater acceptance of the use of AR on smartphones in a private context, with a certain rejection of its use in a physical shop. In general, there is a rejection of the use of AR mirrors in the physical shop due to the lack of privacy during use. However, these problems could be solved by choosing the right location in the shop. Also, the presence of staff who can help and advise during use seems to be of interest. Concerning to the use of Virtual Try-on on mobile phones, there is a certain mistrust that the shades shown are not the real ones, which is the main problem identified.

Keywords Augmented reality · Virtual mirror · Perceptions · Generation X · Generation Z

1 Introduction

Over the last decade, technology and marketing have become inseparable partners. Technological advances have been successfully integrated into marketing strategies given rise to numerous advantages for both customers and enterprises.

S. Barta (✉) · R. Gurrea · C. Flavián

Faculty of Economy and Business, Universidad de Zaragoza, Gran Vía 2, Zaragoza, Spain
e-mail: sbarta@unizar.es

The evolution of AR is allowing new uses and applications to emerge, some of which may be very interesting for a certain sector. AR can improve consumer knowledge about the products offered, while they enjoy the process of trying the items on themselves effortlessly through their phone or smart device. Many industries that focus on providing value for consumers such as retailing, tourism, fashion, entertainment, automotive, services or education may benefit from this technology to deliver enhanced experiences to their consumers. The beauty industry is currently undergoing a profound digitalization process with a 40% global worth in online cosmetic sales during 2020, accounting for 22% of the sales in the industry [1]. The current COVID-19 crisis has sped up the adoption of AR since more people were online shopping and many beauty stores do not allow trying products for sanitary reasons. In this sense, this crisis has highlighted how AR can be a realistic tool to facilitate the shopping experience from home. Before the pandemic just 8% of the US retailers wanted to offer Virtual Try-on to their customers, now the percentage has raised to 20% [1].

Academic research has recently focused attention on the study of the use of AR in consumer behaviour [2, 3]. A lot of research has followed a quantitative methodology, conducting many experimental designs. However, the scarcity of exploratory studies carried out is striking [4], particularly given the novelty of the topic and the need to explore in-depth the perceptions and causes of user behaviour. For this reason, this research aims to contribute to the knowledge of the perceptions and factors that motivate the use of AR both online and in the physical shop, by conducting two focus groups. It also sheds light on in which shopping situations the use of different types of AR may be more useful. To this end, the topic is addressed with different age groups, and some differences between generation X and generation Z can be identified.

2 How to Implement Augmented Reality in the Shopping Decision Process

Some steps should be considered for the successful implementation of AR campaigns in the context of marketing [5].

First, it is necessary to define target audiences and communications objectives before starting to design the campaigns. The campaign can follow completely different patterns depending on to whom it is directed. As an example, generation Z and Millennials are more likely to be keen on technology than “Boomers” or elderly generations, so it is easier for them to participate in campaigns that involve a more complex use of technology [6].

Secondly, it has to be established how the AR filter will be activated. Marketers need to decide when and how the AR layer will be activated. When the user is given the option to activate the AR layer, they feel more in control of the situation and are more likely to have a positive response to it. However, when they do not have to take

any deliberate action, the time and space are precisely selected by the marketer and there is a reluctance to engage at the beginning [7].

Thirdly, the information and content to be provided by the AR filter should be regulated according to its objectives and content. Next, designers will establish the level of interactivity to give users. Interactivity refers to the extent to which users can participate in modifying the form and content of a mediated environment in real-time [8]. This can provide fun and playfulness, which have been shown to have a positive impact on consumer satisfaction. Control, responsibility, personalization and perceived connectivity are necessary characteristics to achieve interactivity [9]. However, when one of the goals of your AR campaign is to provide very low-cost information or there is a likelihood that people will use it maliciously (e.g. showing hate speech), it may be necessary to have strong control over the content and actions available [10].

Finally, establish how the AR information layer will be integrated with the specific social and physical contexts. The last step involves selecting how the AR content will be integrated with the passive elements: the bystanders and the background. In this sense, research has distinguished between augmented reality and mixed reality [11]. Depending on the type of experience a particular brand intends to offer, it will be crucial to decide how much of the context content will be integrated. Higher integration of a certain layer in the context is more resource-intensive in terms of planning, investment and technical capacity requirements. Furthermore, if the content is to be presented in a public place, bystanders have to be taken into account as elements of interaction that could interrupt and ruin the experience (voluntarily or involuntarily). These threats should always be kept in mind and avoided whenever possible. The goal is to engage the consumer at the highest level, so all of these considerations should be assessed and taken into account.

3 Method

After looking at the different technical steps to carry out the correct implementation of AR, through qualitative methodologies we aim to find out the general impressions of this tool, and the perceptions and opinions of the different types of AR that are available for the purchase of beauty products.

For this purpose, we carried out qualitative research (two focus groups). Qualitative methodologies have been deemed appropriate due to their ability to provide a deep understanding of consumers' perceptions and feelings [12]. In this kind of discussion, participants can give their opinion and express their feelings freely. This fact, enhance the interaction between participant and many different points of view arise enriching the quality of the information [13].

The interaction between the researcher and the participants is the origin of valuable data. These data are subjective; however, once interpreted can have an extraordinary value for the research [14]. The subjective experiences described by participants help us to understand their reality and behaviours. Observing participants' interactions,

corporal languages and ways of communicating in a specific context give rise to a lot of interpretable data that would be unmeasurable using quantitative methods.

3.1 Focus Group Design

We conducted two focus groups with consumers who usually buy online. One focus group was formed of six women between 18 and 21 years old (digital natives), and the other of six women between 41 and 57 years old (digital immigrants). Small qualitative sample sizes are considered appropriate for the investigation of emerging concepts [15]. The recruitment process started with postgraduate students over 18 years of age and continued via snowballing.

Generally, the number of participants ranges from four up to twelve; in this case, we opted for small groups of 6 participants to create a more comfortable environment that favours the deepening on more specific issues [16]. Group's composition should have a certain degree of homogeneity to avoid great contrasts, but at the same time, they must be heterogeneous enough to encourage debate and different criteria. Thus, we selected groups with similar characteristics in terms of gender, age and buying habits. Participants should be females of similar ages with an interest in purchasing beauty products, but they came from different backgrounds and had no close personal relationships with each other. The respondents were advised they were participating in a research project to better understand their shopping experiences with AR. The focus group sessions lasted between 60 and 90 min and were audio-recorded. Before the data analysis, the research team checked the transcripts for accuracy.

The sessions took place during the first semester of 2021, and due to the current COVID-19 sanitary restrictions, the focus groups had to take place using Google Meet online conference. However, to increase engagement and spontaneity by having a close feeling all the participants kept their microphones and cameras on. The two groups of people that can use AR technology have the following characteristics [17]:

- (1) Digital natives that have grown in close contact with technologies such as Millennials and generation Z, born after the 80 s.
- (2) Digital immigrants have had to adapt to new technologies, such as Boomers and generation X.

During the focus groups, respondents first described their perceptions and experiences of AR as a Virtual Try-on tool in general and then were presented with an AR as a Virtual Try-on technology to use. Later, it was shown a video of AR in-store, with the use of virtual mirrors. Consequently, a series of short easy questions were asked by the moderator to "break the ice" creating a more comfortable atmosphere. Participants were then asked to try-on virtual make-up on their smartphones and then asked about their experiences using it. Finally, a video was shown of a user using a virtual mirror in the shop, which allowed her to try-on different combinations of make-up on her face and provides personalized recommendations based on her skin tone and other factors. Table 1 shows the sample profile of participants.

Table 1 Sample profile

Participant	Age	Profession	Group
1	21	Business student	Generation Z
2	23	Journalist	Generation Z
3	18	Chemist student	Generation Z
4	21	Business student	Generation Z
5	20	Architecture student	Generation Z
6	21	Management student	Generation Z
7	57	Doctor	Generation X
8	52	Tourism teacher	Generation X
9	56	Technology teacher	Generation X
10	41	User experience designer	Generation X
11	49	Teacher	Generation X

4 Results

Our findings reveal three main themes concerning the perceptions and use of AR in different contexts. Below, we outline each theme.

4.1 *Augmented Reality Perception*

All the participants agreed that AR can help them decide which products to buy, especially when there is a great range of colours that look similar in the photographs. Participants stated that tones that looked the same in the webpage photographs looked completely different once they had tried them out with the AR tool: “It is especially helpful when two colours seem very much alike, then you try them out and you see a clearer distinction” (P1, 21).

A participant in the aged focus group even stated that the experience was similar to going to the store: “it is almost as if you were going to the store to choose the product!” (P8, 52) while another noted that this experience could be even better cause you see the result directly in your face: “in the store, they test it on your hand, not on your lips, so you can see how it suits you and your skin” (P10, 41).

In addition, all of them agreed that they will be opened to try many different products since they can do it effortlessly from their homes with just a few clicks without having to remove the make-up afterwards and making the process much faster than physically trying them out: “I think you try many more products online, for convenience and hygiene reasons” (P2, 23). Moreover, they will try products they had never thought about trying just because they have the possibility in hand, and they feel anonymous using the tool. Some participants will do it just for fun because they find the process of online trying entertaining and fun: “It makes the shopping experience more interactive. You are more entertained, you try things out, you see

how it looks on your face. It's a more fun process, it's something innovative, it's a new experience. It motivates you to try products. Now if you buy by just looking at photos, it is boring" (P4, 21). This could become a problem, due to the large number of alternatives evaluated, as previous research has shown [4].

Therefore, we can say that the interactive of the application makes the client experience an entertaining moment enhancing his purchasing experience and that they are more likely to try new products, since they feel in a safe environment, and this can help them make their final decision.

The younger generation remarked that sometimes they do not like asking for help from the staff of the stores while trying many different products because they feel like they are making them lose their time if in the end up not buying anything, so having this technology in hand could also help them to solve this issue: "If in the end, you are not going to buy anything after having tried on many things, you feel certain remorse. But if someone is advising you are also embarrassed to say that you do not want it. In the application, you try it and that's it" (P3, 18).

4.2 In-Sore Augmented Reality Mirrors

However, participants did not seem to like the idea of trying this AR technology in the stores, they stated that one of the aspects they like about the AR application is that you can try it from your home anonymously without anyone watching you. If this technology is carried out to the store where other people can see them trying things out on a screen, they won't feel motivated to try as many different products.

Regarding this matter, the Millennial and generation Z group added that having this technology in the stores will not suppose an attraction to them since they are used to seeing this kind of filter daily on social media so seeing it in a store will not impact them as much: "I do not think it is very useful to have the same tool that you already have on your phone and you can do it whenever you want at home" (P2, 23); "in the case of makeup, it is not so new. you may not have seen it in stores, but you can already do it on your cell phone..." (P1, 21).

Also, there is a certain rejection due to factors such as use in a public place, such as a shop: "I would never use the mirror in the shop. If I want to try it out virtually, I prefer to do it at home and then go to the shop. If I have any doubts, I will ask in the shop, but I will go with clear ideas. Maybe I would use it for the novelty, but I do not think I would use it for real. I would prefer to do it from home online and then if I have any doubts I can go to the shop. The other option would also be to buy directly online or go to the shop to see the product physically." (P6, 21); "I think I would also use the AR at home because I would be a bit embarrassed to go to the shop and be using the mirror. It is a question of privacy. I would not want other people to see me trying it on. But I do think there will be more. Now in the shops, because of covid, the tester has been removed" (P1, 21).

For generation X, the aspect of privacy in use is also seen as a factor of rejection for the use of AR in the physical shop: "I do not think the virtual mirror in the shop

makes sense. One of the good things about trying things on virtually is the anonymity, not that you are in the shop, and everyone is looking at you, I would rather try it on online". (P11, 49). However, some people are in favour of its implementation in physical shops, subject to certain conditions, such as the availability of staff to help you or that the mirror is located in a place that allows a certain degree of privacy, that not all shoppers can observe you while you are using it: "I like to be recommended and I find it very important to try the products, especially the textures and the scent. I think it's great to try the products on the mirror in the shop if someone is advising me which ones to try and that there is some privacy while you are using them" (P7, 57).

4.3 Augmented Reality Drawbacks

The biggest drawback with AR for participants is that through this technology you can see how products look in your face; however, you cannot experience or feel the textures, smells, creaminess or flavours of the products. These properties are very important and special when choosing make-up since not all products perform equally in everybody. This is shown in statements like: "indeed, you do not try the textures and I also think that is important and the creaminess is also important, even if give you the information later, you have to see how you feel" (P7, 57).

Another issue is that participants did not completely rely on the veracity of this application, and therefore, they questioned if what they were seeing on the screen was 100% true and feared the possibility of it being different in real life. The colours in the photographs being very different to the ones seen in the simulation seemed to raise more doubts about what the real colour is if the one in the photograph or the one in AR. A participant stated: "I don't know if they reflect exactly the colours that they are. Down below you saw some colours and then when you put them on, they were very different. So, that makes me a little bit doubtful" (P6, 21).

Nevertheless, participants of all ages agreed on the fact that this can help them make their final decision or at least reduce the number of alternatives they have in mind. In the case of having tried out AR and still having doubts, the older group would go to the store to try the products and maybe would ask the professionals their opinions. On the other hand, the younger group would look for online reviews (YouTube videos, comments...), friend's opinions or even trust influencers opinions about certain products and finally if the doubts persisted they would go to the stores to try to reaffirm their decision: "I would also send it to my friends. Anything that is augmented reality is much easier to share, especially on mobile" (P2, 23); "Now there are beauty influencers who recommend products and I rely quite a lot on the opinion they have" (P3, 18).

5 Discussion and Conclusions

This technology has huge implications in the field of e-commerce since the consumer is not able to see or try the products until they are delivered to their homes. This may be a huge problem for beauty products buyers since colours and looks play a very important role in their buying decision.

One of the principal reasons for not buying make-up online is the distrust in the shopping decision process. For most of them, it is important to feel and try the product before buying it. Finding a product that fits with their characteristics (skin type, facial shape, skin tone) is crucial, especially if they want to use it for a special occasion. In the traditional online process, potential buyers cannot see how the product looks on them and this is what ultimately causes them to avoid the online process or, if they do proceed with the purchase, to experience some discomfort due to the uncertainty of having chosen the wrong product [18].

AR has been proved to help participants in the decision process. Being able to see how the product may look into their faces eliminates one of the unknowns. This process can be particularly beneficial when there are a lot of options that at first sight look very similar in the photographs. In addition, this technology encourages buyers to try-on many different things, even those who are out of their comfort zone and may end up buying something they had never thought about. The experience of trying different things seems to be fun and enjoyable for shoppers because it does not require as much effort and time as physically trying them, it is more hygienic, and they can do it in a private environment without being observed or feeling judged by others.

Younger generations are the drivers of the industry with the biggest expenditure in this area [19]. In this sense, AR marketing in social media (e.g. Instagram filters) can be useful to attract these young consumers since they will probably engage with the filters without even realizing just to have fun and share the content in their accounts. If they have doubts about the product, they will be more likely to look for information online, reading reviews on social media or listening to influencers opinions.

On the other hand, Generation X will still prefer going to the store to try the products out or been given the opinion from the professionals. This technology can help them try more products and reduce the number of options they have in mind but most of them will still prefer going to the store in the case of being buying something they have never tested.

An important drawback of this technology is the lack of trust in the veracity of the results. This technology is relatively new, and consumers do not fully rely on his ability to show a precise image. However, this technology is going to be more and more common and if people have good experiences buying with it, they will gradually start to trust it. Moreover, despite the benefits, AR has shown a realistic image of how the products will look in the consumers face, consumers are still not able to feel the products, touch them, and check their texture, creaminess.

Finally, given the differences observed between the generations, AR technology should be mainly targeted at "younger generations" that are the ones that are already

adapted to this kind of technology and believe more in the online process. AR can be the solution to a more hygienic and fast way of testing cosmetics, so even if the consumer does not complete the online process, this technology will also allow them to have a clearer idea of what they want and streamline offline purchasing process.

References

1. L'Oreal. <https://www.loreal-finance.com/en/annual-report-2020/cosmetics-market-2-1-0/>. Last accessed 1 Sept 2021
2. Hilken, T., Keeling, D.I., de Ruyter, K., Mahr, D., Chylinski, M.: Seeing eye to eye: social augmented reality and shared decision making in the marketplace. *J. Acad. Mark. Sci.* **48**(2), 143–164 (2020)
3. Javornik, A.: Augmented reality: research agenda for studying the impact of its media characteristics on consumer behaviour. *J. Retail. Consum. Serv.* **30**, 252–261 (2016)
4. Romano, B., Sands, S., Pallant, J.I.: Augmented reality and the customer journey: an exploratory study. *Australas. Mark. J.* **29**(4), 354–363 (2021)
5. Scholz, J., Smith, A.N.: Augmented reality: designing immersive experiences that maximize consumer engagement. *Bus. Horiz.* **2**, 149–161 (2016)
6. Turner, A.: Generation Z: technology and social interest. *J. Individ. Psychol.* **71**(2), 103–113 (2016)
7. Irshad, S., Awang Rambli, D. R.: Multi-layered mobile augmented reality framework for positive user experience. In: *Proceedings of the 2nd International Conference in HCI and UX Indonesia*, pp. 21–26. (2016)
8. Steuer, J.: Defining virtual reality: dimensions determining telepresence. *J. Commun.* **42**(4), 73–93 (1992)
9. Lee, T.: The impact of perceptions of interactivity on customer trust and transaction intentions in mobile commerce. *J. Electron. Commer. Res.* **6**(3), 165 (2005)
10. Thompson, C.J., Rindfleisch, A., Arsel, Z.: Emotional branding and the strategic value of the doppelgänger brand image. *J. Mark.* **70**(1), 50–64 (2006)
11. Flavián, C., Ibáñez-Sánchez, S., Orús, C.: The impact of virtual, augmented and mixed reality technologies on the customer experience. *J. Bus. Res.* **100**, 547–560 (2019)
12. Hammersley, M.: *Methodology: Who Needs It?* Sage, London (2010)
13. Fernández-Nogales, Á.: *Investigación y técnicas de mercado*. Esic Editorial, Madrid (2004)
14. Creswell, J.W.: *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage publications, Thousand Oaks (1998)
15. Van Esch, P., Van Esch, L.: justification of a qualitative methodology to investigate the emerging concept: the dimensions of religion as underpinning constructs for mass media social marketing campaigns. *J. Bus. Theory Pract.* **1**(2), 214–243 (2013)
16. Krueger, R. A., Casey, M.A.: *Focus Groups: A Practical Guide for Applied Research*. Sage Publications (2014)
17. Kesharwani, A.: Do (how) digital natives adopt a new technology differently than digital immigrants? A longitudinal study. *Inf. Manag.* 103–170 (2020)
18. Pang, J., Keh, H.T., Li, X., Maheswaran, D.: “Every coin has two sides”: the effects of dialectical thinking and attitudinal ambivalence on psychological discomfort and consumer choice. *J. Consum. Psychol.* **27**(2), 218–230 (2017)
19. Khan, H.: Role of augmented reality in influencing purchase intention among millenials. *J. Manag.* **6**(6), 37–46 (2019)

Place Attachment Through Virtual Reality: A Comparative Study in Douro Region (Northern Portugal) with Video and ‘Real’ Visit



Nieves Losada , Filipa Jorge , Mário Sérgio Teixeira , Nuno Sousa , Miguel Melo , and Maximino Bessa 

Abstract Immersive technologies, such as virtual reality, could be effective marketing tools for destination marketing, namely in creating place attachment prior experience the destination. Place attachment plays a significant role in behavioural intentions to visit and to recommend a destination. However, place attachment research is relatively new in the tourism context. This study seeks to empirically examine the effectiveness of Virtual Reality in creating place attachment to destinations exploring the changes in the place attachment after two moments. First, after watching a video and after having an experience in the Virtual Reality environment. Second, after the experience in the Virtual Reality environment and after the ‘real’ visit to a representative viewpoint in Douro region. Students belonging to Gen Z were sampled. Findings reflect that Virtual Reality has potentialities for marketing destinations.

Keywords Virtual reality · Place attachment · Douro region

1 Introduction

Immersive technologies, i.e. Virtual Reality, could impact the tourism industry in the coming years if people were readably able to experience locations and attractions from their own homes and from anywhere in the world [1, 2]. Furthermore, from

N. Losada (✉) · M. S. Teixeira

University of Trás-os-Montes e Alto Douro (UTAD), Quinta de Prados, 5001-801 Vila Real, Portugal

e-mail: nlosada@utad.pt

CETRAD Research Unit, University of Trás-os-Montes e Alto Douro, (CETRAD), Quinta de Prados, 5000-801 Vila Real, Portugal

F. Jorge · N. Sousa · M. Melo · M. Bessa

Institute for Systems and Computer Engineering, Technology and Science, INESC TEC, Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

N. Sousa

Faculdade de Ciências Empresariais e Turismo, Universidade de Vigo, Ourense, Spain

a marketing perspective, Virtual Reality can enhance visitor experiences, not only at tourism destinations, but also prior to visiting them [3–5]. Also, Virtual Reality technology could act as a complement or substitute of physical journeys or ‘real’ holidays for some travellers, i.e. for a low-income level or reduced mobility [1, 6].

Virtual Reality adds new scopes to branding and marketing tourism destinations and can be used to complement or even supplant traditional promotional tools such as brochures or video, among others [1, 7, 8]. Research about adoption of Virtual Reality for destination marketing purposes has increased during the last years [9–13]. However, the literature on Virtual Reality applications in tourism is predominantly conceptual [14].

The main fields of interest in Virtual Reality applied to tourism marketing and destination promotion in the last years were destination image [10, 12, 15], and intention to visit/revisit and to recommend a destination [2, 16–23]. Some authors indicated that Virtual Reality can also benefit place attachment in tourism contexts [2, 24, 25]. Place attachment plays a significant role in behavioural intentions to visit and to recommend a destination. However, the potential of Virtual Reality in creating place attachment to destinations has been neglected.

The objective of this study is to deepen in the potential of Virtual Reality to be used as a promotional marketing tool for development place attachment prior to visiting a tourism destination in Gen Z by exploring the changes in place attachment across different moments: after watching a video of the touristic destination site, after a Virtual Reality experience depicting the touristic destination site, and after the visit to a ‘real’ site, a representative viewpoint in Douro region.

2 Theoretical Framework

Virtual Reality enhances perceived destination image compared with traditional promotional tools, such as video and websites [10, 12, 15]. Moreover, Virtual Reality is considered a powerful marketing tool because it impacts tourists’ intention to visit/revisit and to recommend a destination [2, 16–23].

As noted by [26] people develop attachments with environments and places in response to the cognitive and affective components of destination image, and subsequently intention to visit/revisit or recommend the destination [27]. Experiencing a place can result in the formation of emotional connections to a place, and motivate people to return or want to return to that place [25]. However, in spite of its importance for the tourism industry, place attachment is a concept that has been widely explored in psychological, geographical, and recreational fields, but only in recent times in the tourism context [25, 28].

Place attachment has been conceptualized by [29: 275] as “an affective bond or link between people and specific places”. It consists of two dimensions: (1) place dependence, which refers to a functional attachment to a place, this is, “how well a setting serves goal achievement given an existing range of alternatives” [30: 234] and (2) place identity, which refers to a symbolic, emotional, or affective attachment to a

place [31]. In the recreational field, destination attachment has been conceptualized as “a relational construct consisting of the cognitive and emotional connection visitors develop toward a destination” [32: 196].

In tourism, place attachment is often measured as the length of stay. Thus is, the greater the amount of time spent in a place, the stronger attachment levels will be developed [28, 33]. Previous studies have demonstrated that people feel attached to places they lived before or visited as a tourism destination. But, others support that people could develop an attachment to places not experienced before [34]. Individuals can develop place attachment to tourism destination prior to visit it stimulated by media or storytelling (in oral, written or visual form), among others [34].

Although place attachment in the tourism context has been explored mainly after the place visit, recent literature suggests that visitors can develop an attachment to a destination before experiencing it physically [33–36].

A study conducted in Xi’an city airport by [35] showed that place attachment (dependence dimension) mediated the relationship between destination image (cognitive dimension) and perceived service quality prior to visiting a destination.

Another study setting in the Swiss Alps and using ‘Heidi’, a famous literary and television series persona, as a “virtual agent” revealed that place attachment and motivation are keys determinants of intention to visit the destination [34].

One more study dealing with South Korean tourists’ perceptions of Vietnam showed that destination image, both cognitive and affective dimensions, was significant in predicting place attachment before travel [33].

Finally, the study of one more author [36] examined pre-trip attachment formation via mobile technology and social media. Results revealed that text and video on a destination’s Facebook page influence mental imagery and subsequently attachment to that destination.

Apart from that, it has been pointed out that Virtual Reality can benefit place attachment in tourism contexts [2, 24, 25]. In a recent qualitative study with a small sample, researchers [24] found that virtual reality positively influences place attachment by enhancing spatial cognition and positive feeling about the destination. However, research on place attachment in Virtual Reality remains scarce [18, 24, 25].

Examining the shift in place attachment after the virtual tourism experience and before the ‘real’ visit, by comparing with video and site visit, will provide more valuable insights about the efficacy of Virtual Reality for destination marketing purposes.

3 Methodology

The study setting considered was São Leonardo de Galafura viewpoint in the Alto Douro Vinhateiro (Alto Douro Wine Region) world heritage site, Fig. 1. Douro region (NUT III) is a rural area and one of the oldest demarcated wine regions in the world (since 1756), where the world-renowned port wine is produced. It occupies an area of 4100 km², around 19% of the total landmass of Northern Portugal [37]. The number

Fig. 1 Screenshot of the VR experience developed by the research team illustrating the teleport feature



of visitors in the Douro region has increased during the last years due to Portugal's tourism boom [38].

The Virtual Reality experience was designed to be as close to reality as possible. The research team adopted photogrammetry techniques to recreate a hyper-realistic virtual replica of the tourism destination. Participants were allowed to explore the touristic site in the Virtual Reality environment by using real walking within the tracked area (approx. 3.5 m × 3.5 m) or by teleporting using the VIVE remote.

The touristic site allowed to be explored in the Virtual Reality environment was delimited to be coincident with the area that participants could explore in the visit to the real location.

For delivering the Virtual Reality experience, a computer equipped with an Intel i7-6700K and an NVIDIA GeForce GTX 1080 graphics card were used. The visual stimulus was delivered using the HTC Vive setup, and the audio was delivered via Bose QuietComfort 25 headphones with active noise cancellation.

The items utilized in this study “The destination means a lot to me”, “I identify strongly with this destination”, “I feel no commitment to this destination”, “The Douro is the best place for the things I like to do”, “For the activities I enjoy most, the settings and facilities provided by Douro are the best”, “For what I like to do, I could not imagine anything better than the settings and facilities provided by Douro”, “I am very attached to this destination”, are based in [28, 31, 39, 40]. Items were rated on a seven-point Likert scale ranging from ‘strongly disagree’ (1) to ‘strongly agree’ (7).

A pretest was conducted before data collection in order to validate our survey and the experiment procedure, in particular, the comparison process of the above-mentioned variables regarding video, Virtual Reality experience, and actual visit.

This study procedure is composed of two steps, Fig. 2. First, participants arrived at the laboratory at a pre-fixed hour. They came into a room to experiment São



Fig. 2 Experiment procedure

Leonardo de Galafura in a Virtual Reality scene for five minutes. After that, they were moved to another room to experience the same tourist destination in virtual reality, also for five minutes. Following this first step, they were asked to complete the first questionnaire (Q1) in a separate room. This experiment was intentionally designed to participants do not cross during the procedures in Laboratory, in order to not influence other participants' perceptions. Participants were exposed to the video, the Virtual Reality experience, and the 'real' visit for the same amount of time (five minutes) in order to ensure consistency. Second, after joining a small group, participants were conducted to the 'real' site, São Leonardo de Galafura viewpoint, by car or minibus and were encouraged to explore the site for five minutes. Then, participants were asked to rate the same items included in Q1 in order to compare 'pre-travel (after video visualization and Virtual Reality experience)' and 'in situ' responses (on Q2). Finally, participants returned to the laboratory. Finally, participants travelled from the touristic point to laboratory.

The sample was selected with a purposive sampling approach between university students. The only criteria were that participants belong to Generation Z. The sample is considered adequate 'as the comfort and confidence of young students using technology would be less likely to act as a confounding variable', as [15: 7] indicate.

To identify significant differences in participants' responses, relatively to place attachment, two paired sample t-tests were conducted via SPSS 25.0 statistics software. The first paired sample t-test was conducted in the first moment, after the visualization of video and after Virtual Reality experience (Q1), in order to identify significant differences between the promotional tools surveyed. The second paired sample t-test was conducted after the Virtual Reality experience and after the actual visit (Q2) to identify significant differences between Virtual Reality experience and "in situ" experience. The third paired sample t-test was conducted between the deltas (Δ) of the two moments, in order to identify significant differences in the total impact of Virtual Reality experience, and by extension, the promotional capacity of Virtual Reality, respectively to the video and "in situ" experience.

4 Results

From a total of 200 questionnaires distributed, 192 valid questionnaires were returned. The complexity of the study design has favoured this high response rate, 96%, since all the phases of the experiment and data collecting were concentrated on the same day. Within the final sample, 64.1% were national students from different locations of the country. The median age was 20.67 years old, being 24 years old for international students and 19 years old for national students. Most of the participants were females (60.4%). The majority of respondents were attending to a bachelor degree (67.7%), Master (29.7%) or Ph.D. (2.1%); namely in the fields of Management, Economics and Tourism (42.2%), Technology (21.9%) and Others (35.9%) Table 1.

Table 1 Sociodemographics

	\bar{X} or %
<i>Origin</i>	64.1%
Portugal	35.9%
Other nationalities	
<i>Age</i>	20.6
<i>Gender</i>	39.6%
Male	60.4%
Female	
<i>Study cycle</i>	67.7%
Bachelor	29.7%
Master	2.1%
Ph.D.	
<i>Course</i>	42.2%
Management, economics or tourism	21.9%
Technology	35.9%
Others	

The first paired sample *t*-test analysis was conducted to identify whether any significant differences exist in Douro destination place attachment rating after watching the video and after Virtual Reality experience (Q1). As exposed in Table 1, all items had a statistically significant positive change. Therefore, Virtual Reality has demonstrated positive effects when comparing with traditional media (in this case video), not only in relation with perceived destination image, as shown by [10, 12, 15], but also in creating place attachment to tourism destination.

The second paired sample *t*-test analysis was applied to identify whether any significant differences exist in Douro destination place attachment rating after Virtual Reality experience and after the site visit (Q2). As presented in Table 2, all items had a statistically significant positive change. Hence, in spite of the power of Virtual Reality to impact positively in perceptions about a destination, the ‘real’ visit to a representative viewpoint in Douro region has more powerful in creating place attachment than the immersive experience. So, it doesn’t seem likely that Virtual Reality can substitute the real experience, contrary to what was pointed out by some studies [1, 6].

5 Discussion and Conclusion

Virtual Reality technology can enhance visitor experiences complementing or supplanting physical journeys for some travellers [1, 2, 6]. Virtual Reality could be an effective promotional tool in the previsit phase because it can develop place attachment to tourism destinations [2, 24, 25].

The objective of this study was to empirically examine the effectiveness of Virtual Reality for tourism promotion purposes prior to visiting a tourism destination, exploring the changes in place attachment after watching a video, after Virtual

Table 2 T-test sample pairs first moment (Q1) and second moment (Q2) for place attachment

	First moment (Q1)			Second moment (Q2)		
	Video (mean)	VR (mean)	<i>t</i> -value	VR (mean)	'Real' visit (mean)	<i>t</i> -value
The destination means a lot to me	4.07	4.86	-9.57 ^a	4.43	5.17	-8.83 ^a
I identify strongly with this destination	4.2	4.99	-8.55 ^a	4.59	4.23	-7.91 ^a
I feel no commitment to this destination	3.45	3.7	-2.4 ^c	3.7	3.9	-2.6 ^c
The Douro is the best place for the things I like to do	3.82	4.2	-5.57 ^a	4.14	4.63	-7.57 ^a
For the activities I enjoy most, the settings and facilities provided by Douro are the best	3.83	4.28	-5.79 ^a	4.24	4.66	-5.87 ^a
For what I like to do, I could not imagine anything better than the settings and facilities provided by Douro	3.78	4.15	-5.22 ^a	4.08	4.53	-7.03 ^a
I am very attached to this destination	3.99	4.63	-7.21 ^a	4.38	5.03	-7.7 ^a

"a" is " $p < 0.001$ "

"c" is " $p < 0.05$ "

Reality experience and after the ‘real’ visit to a representative viewpoint in Douro region.

Results indicate positive significant differences in all items of place attachment in the first moment (Q1), after watching the video and after Virtual Reality experience, and in the second moment (Q2), after Virtual Reality experience and after the ‘real’ visit. According to previous studies [25, 36] visitors can develop place attachment through virtual worlds before having visited the destination.

Then, Virtual Reality increases tourists’ responses compared with other marketing tools, namely video, as suggested by previous studies [10, 12, 15], and it could be an effective marketing tool in the pre-visit phase, developing place attachment to tourism destination.

Moreover, the results show that ‘real’ visit has a more powerful capacity to create place attachment to destinations compared with Virtual Reality. So, Virtual Reality experience could act better as a promotional tool in creating place attachment to destinations compared with video. However, place attachment originated by ‘real’ visit is greater, compared with Virtual Reality experience.

The major limitation of this study is methodological. The sample was selected with a purposive sample approach, given the complexity of the experience. Also, all the participants belong to Generation Z. Therefore, it’s mandatory to extend this study to other generations to obtain differences between different cohorts.

Moreover, differences in the perceived image of destinations “a priori” and “in situ” from the nationality of the participants were suggested previously [41]. Additionally, given the uniqueness of destinations, it’s necessary to replicate this study in other destinations.

References

1. OECD: Tourism, Trends and Policies. In: <https://www.oecd.org/cfe/tourism/oecd-tourism-trends-and-policies-20767773.htm> (2018)
2. Tussyadiah, L.P., Wang, D., Jung, T.H., Dieck, M.C.: Virtual reality, presence, and attitude change: empirical evidence from tourism. *Tour. Manage.* **66**, 140–154 (2018)
3. Bec, A., Moyle, B., Timms, K., Schaffer, V., Skavronskaya, L., Little, C.: Management of immersive heritage tourism experiences: a conceptual model. *Tour. Manage.* **72**, 117–120 (2019)
4. Collins, S.: The effect of virtual reality on destination image: a case study of education abroad programs. Master thesis. Middle Tennessee State University (2019)
5. Marasco, A., Buonincontri, P., Niekerk, M., Orłowski, M., Okumus, F.: Exploring the role of next-generation virtual technologies in destination marketing. *J. Destin. Mark. Manag.* **9**, 138–148 (2018)
6. Guttentag, D.A.: Virtual reality: applications and implications for tourism. *Tour. Manage.* **31**, 637–651 (2010)
7. Gibson, A., O’Rawe, M.: Virtual reality as a travel promotional tool: insights from a consumer travel fair. In: Jung, T., Dieck, M.C. (eds.) *Augmented Reality and Virtual Reality: Empowering Human, Place and Business*. Springer, UK (2018)

8. Li, T., Chen, Y.: Will virtual reality be a double-edged sword? Exploring the moderation effects of the expected enjoyment of a destination on travel intention. *J. Destin. Mark. Manag.* **12**, 15–26 (2019)
9. Cho, Y.H., Wang, Y., Fesenmaier, D.R.: Searching for experiences; the web-based virtual tour in tourism marketing. *J. Travel Tour. Mark.* **12**(4), 1–17 (2002)
10. Griffin, T., Giberson, T., Lee, S.H., Guttentag, D., Kandaurova, M., Sergueeva, K., Dimanche, F.: Virtual reality and implications for destination marketing. In: Proceedings of the 48th Annual Travel and Tourism Research Association (TTRA), International Conference, Quebec City, Canada (2017)
11. Huang, Y.C., Backman, S.J., Backman, K.F., Moore, D.: Exploring user acceptance of 3D virtual worlds in travel and tourism marketing. *Tour. Manage.* **36**, 490–501 (2013)
12. McFee, A., Mayrhofer, T., Baràrtovà, A., Neuhofer, B., Rainoldi, M., & Egger, R.: The effects of virtual reality on destination image formation. In: Pesonen, J., Neidhardt, J. (eds.) *Information and Communication Technologies in Tourism 2019*, Proceedings of the International Conference in Nicosia, Cyprus, January 30–February 1, pp. 107–119 (2019)
13. Rainoldi, M., Driescher, V., Lisnevskaja, A., Zvereva, D., Stavinska, A., Relota, J., Egger, R.: Virtual reality: an innovative tool in destinations' marketing. *The Gaze J. Tour. Hospitality* **9**, 53–68 (2018)
14. Bogicevic, V., Seo, S., Kandampully, J.A., Liu, S.Q., Rudd, N.A.: Virtual reality presence as a preamble of tourism experience: the role of mental imagery. *Tour. Manage.* **74**, 55–64 (2019)
15. Adachi, R., Cramer, E.M., Song, H.: Using virtual reality for tourism marketing: a mediating role of self-presence. *Soc. Sci. J.* (in press 2020)
16. An, S., Choi, Y., Lee, C.: Virtual travel experience and destination marketing: effects of sense and information quality on flow and visit intention. *J. Destination Market. Manag.* **19**, 100492 (2021)
17. Deng, X., Unnava, H.R., Lee, H.: 'Too true to be good?' when virtual reality decreases interest in actual reality. *J. Bus. Res.* **100**, 561–570 (2019)
18. Dieck, D., Dieck, M.C., Jung, T., Moorhouse, N.: Tourists' virtual reality adoption: an exploratory study from Lake District National Park. *Leis. Stud.* **37**(4), 371–383 (2018)
19. Huang, Y.C., Backman, K.F., Backman, S.J., Chang, L.L.: Exploring the implications of virtual reality technology in tourism marketing: an integrated research framework. *Int. J. Tour. Res.* **18**(2), 116–128 (2016)
20. Lee, M., Lee, S.A., Jeong, M., Oh, H.: Quality of virtual reality and its impact on behavioral intention. *Int. J. Hosp. Manag.* **90**, 102595 (2020)
21. Pantano, E., Servidio, R.: An exploratory study of the role of pervasive environments for promotion of tourism destinations. *J. Hosp. Tour. Technol.* **2**(1), 50–65 (2011)
22. Wei, W., Qi, R., Zhang, L.: Effects of virtual reality on theme park visitors' experience and behaviors : a presence perspective. *Tour. Manage.* **71**, 282–293 (2019)
23. Yung, R., Khoo-Lattimore, C., Potter, L.E.: VR the world: experimenting with emotion and presence for tourism marketing. *J. Hosp. Tour. Market.* **46**, 160–171 (2021)
24. Pantelidis, C., tom Diek, M.C., Jung, T., Miller, A.: Exploring tourist experiences of virtual reality in a rural destination: a place attachment theory perspective. *E-Rev. Tour. Res.* 1–5 (2018)
25. Plunkett, D.: On place attachments in virtual worlds. *World Leis. J.* **53**(3), 168–178 (2011)
26. Baloglu, S., McCleary, K.W.: A model of destination image formation. *Ann. Tour. Res.* **26**(4), 868–897 (1999)
27. Loureiro, S.M.C.: The role of the rural tourism experience economy in place attachment and behavioral intentions. *Int. J. Hosp. Manag.* **40**, 1–9 (2014)
28. Kyle, G.T., Absher, J.D., Graefe, A.R.: The moderating role of place attachment on the relationship between attitudes towards fees and spending preferences. *Leis. Sci.* **25**, 33–50 (2003)
29. Hidalgo, M.C., Hernández, B.: Place attachment: conceptual and empirical questions. *J. Environ. Psychol.* **21**, 273–281 (2001)

30. Jorgensen, B.S., Stedman, R.C.: Sense of place as an attitude: Lakeshore Owners' attitudes toward their properties. *J. Environ. Psychol.* **21**(3), 233–248 (2001)
31. Prayag, G., Ryan, C.: Antecedents of tourists' loyalty to Mauritius: the role and influence of destination image, place attachment, personal involvement, and satisfaction. *J. Travel Res.* **51**(3), 342–356 (2012)
32. Morais, D., Lin, C.H.: Why do first-time and repeat visitors patronize a destination? *J. Travel Tour. Mark.* **27**, 193–210 (2010)
33. Kim, S., Styliadis, D., Oh, M.: Is perception of destination image stable or does it fluctuate? A measurement of three points in time. *Int. J. Tour. Res.* **21**, 447–461 (2019)
34. Hosany, S., Buzova, D., Sanz-Blas, S.: The influence of place attachment, ad-evoked positive affect, and motivation on intention to visit: imagination proclivity as a moderator. *J. Travel Res.* **59**(3), 477–495 (2020)
35. Cheng, P., Qiang, W.: Pre-visit destination image's effects on perceived travel service quality: the mediating effects of place attachment. In: 15th International Conference on Service Systems and Service Management (ICSSSM), July, 21–22, China (2018)
36. Plunkett, D.: Examining the impact of media content, emotions, and mental imagery visualization on pre-trip place attachment. Doctoral dissertation. Arizona State University (2013)
37. GEPE: Projecção do Parque Escolar por NUT III a 2013 – Sub-região Douro [Projection of the School Park by NUT III in 2013 - Douro Sub-region]. Faculdade de Letras da Universidade de Coimbra, Coimbra (2011)
38. WTTC: Travel & Tourism: Economic Impact 2018 Portugal (2018)
39. Veasna, S., Wu, W.Y., Huang, C.H.: The impact of destination source credibility on destination satisfaction: the mediating effects of destination attachment and destination image. *Tour. Manag.* **36**, 511–526 (2013)
40. Yuksel, A., Yuksel, F., Bilim, Y.: Destination attachment: effects on customer satisfaction and cognitive, affective and conative loyalty. *Tour. Manag.* **31**, 274–284 (2010)
41. Iordanova, E., Styliadis, D.: International and domestic tourists' 'a priori' and 'in situ' image differences and the impact of direct destination experience on destination image: the case of Linz, Austria. *Curr. Issues Tour.* **22**(8), 982–1005 (2019)

The Use of Virtual Reality as an Advertising Tool for Brand Experience in Peru



Javier Espinoza-Nieves and Francisco Arbaiza

Abstract Recently, the advertising industry has implemented new methods to improve the relationship between brands and their consumers. To understand how to achieve this effect in a positive way, several studies have turned to the examination of advertising actions aimed at boosting brand experience. However, there are few studies that deepen the knowledge of the importance of creating brand experiences through innovative communication technologies such as virtual reality (VR). The present research explores how the use of VR in advertising helps in the generation of brand experiences. A qualitative methodology of phenomenological design was used with advertising professionals in the city of Lima. In addition, the exploratory scope of the analysis followed a descriptive coding process. It was evidenced that the properties of VR, such as immersion, presence, and its sensory capabilities are crucial to promote brand experiences with greater intensity. This category of virtual brand experiences influences brand image, affects purchase intentions and, above all, builds brand loyalty.

Keywords Virtual reality · Advertising · Brand experience · Immersion · Technology

1 Introduction

Advertising is in a stage where creating a connection between a brand and its target audience is a task that involves a great effort for companies. Despite this challenge, marketing and advertising employ several strategies to strengthen this link, and one of them is the creation of brand experience [19]. This concept is understood as the by-product of the consumer's perception in all the contacts develops with the brand [6]. In this regard, one of the new doctrines that have explored the effective use of

J. Espinoza-Nieves · F. Arbaiza (✉)
Peruvian University of Applied Sciences, Prolongación Primavera 2390, Lima 15023, Peru
e-mail: francisco.arbaiza@upc.edu.pe

J. Espinoza-Nieves
e-mail: u201616749@upc.edu.pe

brand experiences is experiential marketing and, specifically, multisensory marketing [5, 28].

This discipline, by its nature, relies on new technologies to carry out advertising actions and, within these advances, virtual reality (VR) stands out [17, 20, 26]. This technology is defined as an immersive and interactive experience in which users are immersed in a specific space artificially created by computerized mechanisms [3, 5, 16]. Moreover, it has become a technological phenomenon whose wearable devices such as smartphones, computers, or special viewers become increasingly advanced and affordable to the common public as the years go by [23, 26, 27].

Therefore, it is possible to establish that VR is a promising advertising platform and a valid promotional alternative for many market categories [14, 16, 18, 27]. This is one of the reasons behind the existence of several studies about the technological potential of VR in the advertising field; however, there are very few that evaluate its ability to produce an effective brand experience with significant effects on the perceptions and consumption behavior of the participating public.

For this reason, the purpose of this research is to review the use of virtual reality in advertising as a tool to enhance brand experience. To understand this objective, the following research question was identified:

PI: In what way does the use of virtual reality in advertising favor the creation of brand experience?

To answer this question, this paper intends to provide two relevant contributions: (i) to review the capabilities and limitations of VR in the creation of advertising content with the objective of brand experience, and (ii) to identify the benefits that the advertising use of VR provides to brands and advertising agencies.

2 Brand Experience in Advertising Through VR

Currently, VR is conceptualized as a medium composed of computerized, suggestive, and interactive digital simulations or representations that detect the participant's position and replace or augment the input from one or more senses, thereby offering the user the sensation of being mentally immersed in a virtual world [4, 7, 11, 27].

Regarding the key features of VR, Barnes [2], addressed the concept of 'presence' as the experience where users feel that they are truly in the synthetic environment presented to them. Similarly, there are two main dimensions that grant the feelings of presence experienced by humans in virtual reality: 'vividness' and 'interactivity' [2, 3]. On the one hand, vividness is the virtue that VR possesses of producing mediated environments that appear sensorially real and, on the other hand, interactivity is the degree to which users of a medium can influence the form and content of the virtual environment presented to them [2, 23].

Another main element in the study of VR characteristics is found in 'immersion', a term related to presence and referring to the way in which the users perceive stimuli from a virtual world, so that they feel that they lose contact with the physical

reality around them [13, 20, 21]. Van Berlo et al., [24], studied ‘presence’ through three other factors interconnected with immersion: ‘involvement’, understood as the state of the participants’ attention on the elements of the VR experience; ‘sensory fidelity’, understood as the coherence of sensory stimulations; and ‘interface quality’, considered as the overall performance of the technology.

Considering all these conditions that VR features, several experts [11, 25], have ventured to claim that it is one of the most promising information technologies for business innovation. Undoubtedly, one of the industries that could be most favored by VR is the advertising industry, in which this technology is considered as the digital advertising platform of the new generation [12, 27]. VR offers marketing a novel, fresh and creative alternative to reach consumers through virtual interaction with products and services [3, 25].

According to Kilic et al. [15], the advanced pace at which virtual reality is developing has created new ways of conducting advertising businesses. Advertising in combination with VR contains a more immersive message, becoming a communicative resource to create greater brand exposure [14]. For this purpose, equipment such as smartphones, PCs, VR visors such as HMDs, among other accessories, are often used for the creation of immersive virtual advertising experiences [20, 26, 27]. Advertising agencies that have experimented with the creation of advertising content through this technology have stated that the communication actions have managed to provoke emotion, joy, and fun in their users so that the level of engagement in their participation has been high [12].

Thus, among the technological tools available for the origination of brand experience in marketing and advertising, VR stands out, being a technology that has been transformed into a means to propitiate more reliable narratives and experiences and that have the power to cause positive changes in the engagement or emotional attachment between brands and consumers [2, 7, 12].

Similarly, VR technology possesses the ability to affect consumer behavior and their perceptions of the products they are regularly exposed to, such that it is an instrument that maximizes the impact on purchase intentions when used as a brand experience tool [3, 5, 11, 17].

As Gauquier et al. [10] state, since the emergence of VR, several brands have taken the initiative to implement it in their advertising strategies, with the purpose of evoking a brand experience that directs consumers to develop brand stimuli in their minds. For these authors, this technology represents the opportunity to convert passive consumers into active participants of a more pronounced experience by involving sensory stimuli and responses. For this reason, consumers who are exposed to this technology tend to enjoy a brand experience in which the sensory dimension is more pronounced [27].

3 Methodology and Procedures

The objective was to develop an introspective approach about the perception that advertising professionals have about VR in the advertising business and its relationship with the development of brand experience. To analyze the way in which the use of VR favors the creation of brand experiences and to expose the limitations and capabilities of this technology in its use in advertising, a qualitative research approach was used.

The analytical approach of this research was developed through a phenomenological design, which is based on the information that the study subjects provide in first person about a phenomenon whose purpose is to answer the research question [9]. In other words, according to Alase [1], this design provides the researcher with the ability to understand the phenomenon through the experience of individuals and their perception for the interpretation of relevant findings.

The selection process of the participants for the study was carried out through intentional sampling, a technique widely used in qualitative research useful to obtain cases abundant in information that is strictly concurrent with the phenomenon studied and that enrich it with key information from participants with characteristics that were pre-established for the analysis [22]. In this research, the participants were advertising professionals from Metropolitan Lima, with at least 7 years of work experience and holding a university or technical degree linked to advertising, communication sciences, or related careers.

The data collection process in this research was carried out through in-depth interviews because the subject of advertising VR is a deep field in which a fluid conversation with the participant is necessary. Therefore, there were ten participants in this analysis, because data saturation was reached after the tenth interview. Indeed, several experts specify the importance of in-depth interviews in qualitative studies of phenomenological design, for their ability to function as a tool to learn about the subjective experiences, feelings, and perceptions of individuals [8].

4 Results

At the end of the descriptive coding procedure with the data collected from the interviews with advertising professionals, four key categories emerged from this research.

4.1 Category 1: VR Advertising Communication Capabilities

This category describes the communicative capabilities of VR when used for advertising campaigns. A common conception was evidenced in the fact that the interviewed advertisers consider VR to be an innovative and novel communication platform in terms of its use in the advertising field. For example, one of the participants expressed it as follows: *“[virtual reality] catches your attention and people are always looking for something new, a different experience, it is something you do not expect, (...) it is the first sensation of doing something different and new, I think it helps advertising”*.

In reference to this, two main capabilities of advertising communication with VR are defined: impact and brand recall. Regarding impact, this concept was explained by many of the participants as the level of impression that stands out in advertising actions, while brand recall was defined as the advantage of VR as a tool to keep in the users' minds the brands executing these communications. Regarding impact, one of the participants verbalized it as follows: *“[virtual reality] what it has in favor is the impact; the interaction, it captures the person in a certain time, it can be 30 s or 5 min, but it captures him/her”*. On the other hand, regarding brand recall, others mentioned: *“I think that through the experience that the brand provides to the user it is a memorable resource. In VR, the person fully commits to that moment”*.

Up to this point, according to the perception of several advertisers, the capabilities explained of VR advertising revealed a conclusive point: it is a technology that turns out to be a more feasible way to communicate attributes or benefits of a brand: *“VR helps many advertising products that are difficult to understand how they work, because with VR they can be explained in a better way, and many times, in advertising, the conventional ways to teach how a product works is not really enjoyable.”* In this way, it is understood that VR stands out for its didactic communicative capabilities, which many agents in the advertising world are willing to take benefit of.

4.2 Category 2: VR Versus Traditional Media

The purpose of this category is to explain the debate among advertising professionals as to whether VR can be considered an advertising medium or simply a tool according to its advantages.

Some participants describe this technology only as a “tool” due to its scarce use and standardization in advertising practice, which is why they consider that it is not the main element in the development of communication campaigns. One of the interviewees expressed it as follows: *“For me it is an advertising tool, a communication media is another concept, it is a vehicle by which you get a message to a probably larger group”*.

On the other hand, several other participants declared VR not only as a tool but also as communication media, because it shares with other traditional media the common

characteristic of being able to spread brand messages. For example, the statement “*I consider that it is a medium, you are having a platform reaching an audience, there is a sender and there is a message that you are communicating that has to do with some action that is aligned to the brand*”, shows that VR is a technology that can be categorized as a non-conventional and media.

4.3 Category 3: Presence and Immersion: Sensory Elements in the VR Brand Experience

This category breaks down the most notable innate advantages of VR advertising: presence and immersion, as well as other sensory elements of this technology that motivate the generation of brand experiences.

Brand experience, for the interviewees, is understood as a brand’s practice of allowing its consumers to feel a key element of its products and services. This sensory nature of experiences is what allows VR to be configured as the ideal advertising tool for executing memorable communications in advertising. One of the participants exemplifies this in the following way: “*If a brand manages to awaken senses in its users [with virtual reality], that make them feel part of the brand in first person, that is something much more immersive and real.*”

It should be noted that immersion is a main feature of VR that facilitates the execution of a lasting brand experience over time: “*Being able to keep you trapped is one of the great advantages (...) It is a complete attention on the brand*”. Therefore, immersion is understood by interviewees as the intense feeling of being encapsulated in a fictitious environment.

Likewise, another main feature of VR that enhances brand experiences is presence, a term that interviewees described as an individual’s sense of feeling present in a place they are not actually in. For example, one participant noted, “*Its [virtual reality’s] advantage is that it transports you to a moment. It’s an instant that takes you out of the present and puts you into a situation.*” Therefore, from the perspective of advertising professionals, virtual reality provides experiential brand experiences that captivate the user’s full attention.

4.4 Category 4: Effect of VR Brand Experience on Purchase Intention

This category explains the effect of brand experiences with VR on purchase intention of brands that have disseminated advertising communication of their products or services through this technology. The participants highlighted the effect of immersive tools such as VR, which lead their users, who are potential customers or consumers of various brands, to become buyers and, more importantly, a loyal audience. One

interviewee explained: “The brand uses it [virtual reality] as a benefit if it can identify well what is that element that customers do not motivate them to the purchase decision”.

According to the interviewees, brand experiences where the platform used is VR can be the final component in advertising actions to make a purchase because this technology makes the user experience the sensation of owning and interacting with the brand’s products and services. On this point, an interviewed advertiser mentions: “I believe that VR shows you what is going to happen with that product or service you are planning to acquire, so it helps you to make decisions”. This sense of vividness is a natural characteristic of an optimal VR experience, which pushes the user to make the effort to purchase the product in the real world.

On the other hand, another advertiser expressed that the ability to influence the purchase decision is the main objective of VR if it is used to generate brand experience, a fact he expressed as follows: “It’s no use for me to visualize or desire the product if I don’t end up making a purchase. [The brand experience] is the sum of actions that seek the bond between the brands and their public”. Thus, it is understood that the goal of VR in advertising, for several professionals in this field, is to ensure that the advertising brands obtain favorable results that translate into sales and brand loyalty.

5 Conclusions

According to the results of the qualitative and exploratory analysis, it is proven that VR is a tool that, in the current advertising environment, is suitable for creating a favorable and lasting brand experience over time. The experiential effect of this technology is possible because of its properties of immersion and presence, which are sensory capabilities of VR that generate intense sensations of reality and whose purpose in advertising is to develop positive feelings towards a brand that enhance purchase intentions.

The ability to generate memorable brand experiences is the maximum potential that VR can achieve, a fact that not only provides a number of remarkable benefits such as impact, brand recall, and positive brand feelings, but it is also able to modify the behavior of current and potential consumers of a brand so that they make a purchase [5, 11, 17].

In addition, it is observed that this technology is an excellent means to build and enhance the image of brands that sponsor at some point this virtual category of advertising actions. Therefore, the results show that VR is useful for building brand-consumer loyalty after its use, a situation that represents the maximum expected benefit in the development of virtual brand experiences, despite the high cost perceived by advertisers.

It is important to highlight that an important limitation of VR is, for the participants, the low reach of this technology compared to other traditional mass media, because it is usually reserved for single advertising actions and in specific spaces with

a small or niche audience. This limitation was not addressed by a convincing amount of previous research on VR advertising and may be because the most advanced studies on this topic were carried out in more developed countries such as the United States, where this technology has more alternatives for massification and application, unlike countries such as Peru, where this platform still has a reduced use in the advertising industry.

However, the advertisers interviewed were able to explain the benefits of immersion and impact and the limitations of scope and costs that fall on this technology, in such a way that a debate is opened as to whether it is a medium that competes with traditional media or if it complements the advertising practice. For this reason, it would be very useful for future research to study the advantages and disadvantages of VR advertising compared to mass media, to define whether it is an appropriate tool to be selected as a complementary medium in mass advertising campaigns, or whether its use should be limited only to specific advertising actions.

References

1. Alase, A.: The interpretative phenomenological analysis (IPA): a guide to a good qualitative research approach. *Int. J. Educ. Literacy Stud.* **5**(2), 9–19 (2017)
2. Barnes, S.: Understanding virtual reality in marketing: nature, implications and potential. *SSRN Electron. J.* (2017)
3. Cowan, K., Ketron, S.: A dual model of product involvement for effective virtual reality: the roles of imagination, co-creation, telepresence, and interactivity. *J. Bus. Res.* **100**, 483–492 (2019)
4. Cuesta, U., Mañas, L.: Integración de la realidad virtual inmersiva en los Grados de Comunicación. *Icono14* **14**(2), 1–21 (2016)
5. Deng, X., Unnava, H.R., Lee, H.: “Too true to be good?” When virtual reality decreases interest in actual reality. *J. Bus. Res.* **100**, 561–570 (2019)
6. Dong, H.K., Song, D.: Can brand experience shorten consumers’ psychological distance toward the brand? The effect of brand experience on consumers’ construal level. *J. Brand Manag.* **26**(3), 255–267 (2019)
7. Flavián, C., Ibáñez-Sánchez, S., Orús, C.: The impact of virtual, augmented and mixed reality technologies on the customer experience. *J. Bus. Res.* **100**, 547–560 (2019)
8. Fusch, P.I., Ness, L.: R: are we there yet? Data saturation in qualitative research. *Qual. Rep.* **20**(9), 1408–1416 (2015)
9. Fuster Guillen, D.E.: Investigación cualitativa: Método fenomenológico hermenéutico. *Propósitos y Representaciones* **7**(1), 201 (2019)
10. Gauquier, L., Brengman, M., Willems, K., Van, H.: Leveraging advertising to a higher dimension: experimental research on the impact of virtual reality on brand personality impressions. *Virtual Reality* **23**, 235–253 (2019)
11. Grudzewski, F., Awdziej, M., Mazurek, G., Piotrowska, K.: Virtual reality in marketing communication—the impact on the message, technology and offer perception—empirical study 1. *Econ. Bus. Rev.* **4**(3), 36–50 (2018)
12. Han, K.-S., Cho, J.-H.: The effect of presence and engagement in HMD (head mounted display) virtual reality advertising on attitude and memory. *J. Korea Convergence Soc.* **10**(6), 139–146 (2019)
13. Hudson, S., Matson-Barkat, S., Pallamin, N., Jegou, G.: With or without you? Interaction and immersion in a virtual reality experience. *J. Bus. Res.* **100**, 459–468 (2019)

14. Jung, J., Yu, J., Seo, Y., Ko, E.: Consumer experiences of virtual reality: insights from VR luxury brand fashion shows. *J. Bus. Res.* **130**, 517–524 (2021)
15. Kilic, K., Saygi, M., Sezer, S.: Exact and heuristic methods for personalized display advertising in virtual reality platforms. *J. Ind. Manag. Optim.* **15**(2), 833–854 (2019)
16. Leung, X.Y., Lyu, J., Bai, B.: A fad or the future? Examining the effectiveness of virtual reality advertising in the hotel industry. *Int. J. Hosp. Manag.* **88** (2019)
17. Loureiro, S.M.C., Guerreiro, J., Eloy, S., Langaro, D., Panchapakesan, P.: Understanding the use of virtual reality in marketing: a text mining-based review. *J. Bus. Res.* **100**, 514–530 (2019)
18. Realidad virtual y realidad aumentada, los nuevos mecanismos que potencian la efectividad del mensaje en la publicidad. <https://expeditiorepositorio.utadeo.edu.co/handle/20.500.12010/8136>. Last accessed 2020/08/30
19. Moreira, A., Fortes, N., Santiago, R.: Influence of sensory stimuli on brand experience, brand equity and purchase intention. *J. Bus. Econ. Manag.* **18**(1), 68–83 (2017)
20. Otegui, J.: La realidad virtual y la realidad aumentada en el proceso de marketing. *Revista de Dirección y Administración de Empresas* **24**, 155–229 (2017)
21. Paíno-Ambrosio, A., Rodríguez-Fidalgo, M.: Propuesta de “géneros periodísticos inmersivos” basados en la realidad virtual y el vídeo en 360°. *Revista Latina De Comunicación Social* (74), 1132–1153 (2019)
22. Palinkas, L.A., Horwitz, S.M., Green, C.A., Wisdom, J.P., Duan, N., Hoagwood, K.: Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm. Policy Ment. Health Ment. Health Serv. Res.* **42**(5), 533–544 (2015)
23. Tussyadiah, I.P., Wang, D., Jung, T.H., Tom Dieck, M.C.: Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tour. Manag.* **66**, 140–154 (2018)
24. van Berlo, Z.M.C., van Reijmersdal, E.A., Smit, E.G., van der Laan, L.N.: Inside advertising: the role of presence in the processing of branded VR content. *Augmented Reality Virtual Reality* 11–22 (2020)
25. Van Kerrebroeck, H., Brengman, M., Willems, K.: When brands come to life: experimental research on the vividness effect of virtual reality in transformational marketing communications. *Virtual Reality* **21**(4), 177–191 (2017)
26. Violante, M.G., Vezzetti, E., Piazzolla, P.: How to design a virtual reality experience that impacts the consumer engagement: the case of the virtual supermarket. *Int. J. Interact. Des. Manuf. (IJIDeM)* **13**(1), 243–262 (2019)
27. Wedel, M., Bigné, E., Zhang, J.: Virtual and augmented reality: advancing research in consumer marketing. *Int. J. Res. Mark.* **37**(3), 443–465 (2020)
28. Wiedmann, K.-P., Labenz, F., Haase, J., Hennigs, N.: The power of experiential marketing: exploring the causal relationships among multisensory marketing, brand experience, customer perceived value and brand strength. *J. Brand Manag.* **25**(2), 101–118 (2017)

Immersive Digital Marketing for Smart Cities Focusing Tourism



Carlos R. Cunha , Luisa Lopes , and Vítor Mendonça 

Abstract The tourism sector has gained increasing importance in the economy of many regions, or even countries, however, the way in which organizations and territories communicate and promote each other, to attract tourists, significantly impacts the success of the sector, and digital marketing strategies have stood out as promising. Thus, in a digital society, in which information and communication technologies are present in the daily life of the general population, especially mobile technologies, it is essential that the tourism sector seeks innovative and differentiating marketing and communication strategies. When considering the technological advances, which have been asserted lately, wireless communication networks with high bandwidth, processing capacity of mobile technologies, Internet of Things, Big Data, Virtual and Augmented Reality, Empowered Smart Cities with different technology; it can be said that the conditions are in place so that new paradigms and marketing strategies can emerge. These technologies allow to communicate with the target audience in a more impactful way and stimulating different senses, that is, immersive. It is in this context that the opportunity and motivation for the development of this research work arises. In this sense, this work highlights the technologies that we consider most relevant, combining them in a conceptual model that is proposed and that we consider enhancing an immersive and innovative digital marketing strategy, for the promotion of touristic destinations.

Keywords Immersive marketing · Smart cities · Model · Tourism

C. R. Cunha (✉)

UNIAG; Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

e-mail: crc@ipb.pt

L. Lopes · V. Mendonça

Instituto Politécnico de Bragança, CiTUR, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

e-mail: luisa@ipb.pt

V. Mendonça

e-mail: mendonca@ipb.pt

1 Introduction

The dissemination of Information and Communication Technologies (ICT) and the democratization of internet access, highlighting the processing capacity of mobile technologies and the increase in the bandwidth of wireless networks, has contributed to organizations seeking to implement strategies for differentiating technology-based marketing.

Digital technology has tremendous benefits, it allows them to capture rich data such as photos, videos, and online messages and social media posts [1], but also, the increased mass use of smartphones will also have to be seen as an opportunity to keep customers connected. As such, technology is becoming a preponderant, if not central, element, pushing organizations to adopt digital marketing strategies.

According to Kotler et al. [2], companies should focus on offering new customer experiences on three different levels: informative, interactive, and immersive. Whenever customers need answers, seek to talk, and surround themselves with sensory experiences, companies must be prepared to offer that.

However, marketing practices such as digital search and advertising, social media interaction, mobile tracking and engagement, online purchase, and in-store shopping experience, are increasingly powered by scalable and intelligent algorithms, with the help of both technology powerhouses such as Google and Amazon and many other smaller Marketing Technology (MarkTech) companies [3].

Similarly, ICT has had a great impact on the tourism sector. Technology has transformed the tourism and leisure sectors over the past twenty years in so many ways [4, 5]. ICTs provide unique opportunities for innovative organizations to redesign tourism products to address individual needs and to satisfy consumer wants, in this sense, the internet technology has revolutionized the promotion and communication functions of tourism [6].

Considering the increase in demand for new forms of tourism, highlighting the growing level of demand from tourists for the enjoyment of differentiating, rewarding, and enriching experiences, the tourism sector has sought to constantly evolve and adapt to market demands.

Among the contributions that technology is making to these sectors, Virtual Reality (VR) is creating immense opportunities for the leisure and tourism industries [4], Especially because it allows the tourist to be more effectively involved with the destination. Virtual reality technology uses a computer to create a more realistic virtual environment, combined with the auxiliary functions of different detection equipment, so that the user is completely immersed in the generated virtual environment, and then interacts with the virtual environment through the man-machine interface so that the user will have a kind of illusion like the illusion that the user is in the real world [7].

Technologies capable of creating immersive environments, such as VR, are tools that can play a very important role in attracting new visitors, due to their ability to provide incredible experiences and maximize the likelihood that the visitor will repeat and recommend the visiting experience [8].

Nevertheless, marketing strategies, especially within the tourism sector, should increasingly be transversal to the various moments, concerned, at first, with helping the tourist to make the choice of destination, secondly, during the period of the tourist experience, supporting and promoting more enjoyment of the tourist experience, and finally, after the visit, allowing the tourist to give feedback and maintain interest in revisiting.

Thus, technologies can, and should, integrate marketing strategies designed to help and accompany tourists throughout the pre-visit phase, during the trip, and in the post-visit phase.

Technological advancements are having a deep impact on the tourism sector, since business management technologies (e.g., mobile technologies and cloud computing, automation and advanced robotics, blockchain, data analytics), to technologies that produce innovative tourism products, services, and experiences (e.g., virtual/augmented reality, Internet of Things), and technologies that assist, understand and connect with markets (e.g., data analytics, cloud computing, and artificial intelligence) [9], since their combination, that is, the integration of different technologies, will foster the development of innovative and differentiating strategies, which are shown to enhance competitive advantages and at the same time represent value for customers.

New technologies make possible the convergence between the physical and digital worlds supported by sensors that collect data arising from mutual interactions between tourists and the environment [5]. Thus, we can highlight that the integration of different technologies allows multisensory marketing strategies to be developed and involve the tourist in immersive experiences, experiences which, in turn, can contribute to getting to know the tourist and, depending on the interactions, their motivations, and interests, and eventually, be able to influence their behavior, which can be understood as immersive marketing.

Considering the efforts produced, in recent years, in the scope of Smart Cities, where, Smart Cities have come out as a major initiative by various governments in making cities more navigable and welcoming to the expected population increase and providing city dwellers a better living experience [10], being that Smart cities act as a ladder for the establishment of smart tourism destinations, which utilize the available technological tools, innovations, and techniques to enable pleasure, and experiences for the tourist and profit for the organizations and the destinations [11], we believe that it is useful that the conceptual model, proposed in this paper, considers the reuse of all the infrastructure set up within the scope of smart cities.

After this initial contextualization, in the next chapters, we intend to frame the perspectives of digital marketing and tourism in an era in which several disruptive technologies prevail, and that will support the proposed conceptual model. Finally, we conclude this paper with a brief reflection on the potential of aggregating technologies, under the vision of the smart city concept, as a lever in the development of new marketing strategies focused on tourism.

2 Perspectives for Digital Marketing

Marketing has faced major changes due to technological developments implemented in everyday life. Businesses around the world have been hit hard by the COVID-19 pandemic, where periods of confinement and restrictions on mobility acted as an accelerator for digitization, for customers and companies [2]. Probably, this behavioral change will last beyond the crisis, and so companies should focus on offering the new customer experience and be prepared. Still, larger companies will eventually have a greater advantage over small and medium-sized companies, as they have a larger budget and knowledge to keep up with and keep pace with the transition to this increasingly digital era [12].

According to Kotler et al. [2], the most recent concept of Marketing 5.0 complies with the application of humanized technology (humanized approach to technologies—by imitation of the human being) in order to create, communicate, fulfill and enhance value throughout the customer's journey. In Marketing 5.0, the main theme is the so-called next tech, that is, a set of technologies that aim to emulate the capabilities of marketers. They include Artificial Intelligence (AI), Natural Language Processing (NLP), sensors, robotics, Augmented Reality (AR), Virtual Reality, Internet of Things (IoT), and blockchain. The mixture of these technologies makes Marketing 5.0 [2]. These new-age technologies will be a crucial part of the future of marketing, and it will be inevitable for companies to adopt them. For instance, the core business of advertising, driven by big data and AI technology, is increasingly becoming intelligent, from consumer insight to advertising information delivery and then to advertising content production. With the incessant progress of cognitive computing technology, it can be forecasted that in the future, intelligent advertising will continue to step forward along the trend of “humanization” according to human needs [13].

In the post-digital period, no matter the generation, consumers have more demands, less time to spend, limited attention, and have easily affected and shorter customer journeys [14]. Furthermore, even if preferences and needs are dynamic, opportunities are temporary, so companies need to be constantly aware of consumer demands since, these change in a short period of time, and the opportunity for companies to satisfy them is even smaller [3].

Specifically, AI refers to a set of programs and algorithms that replace human capabilities, with more accurate and real-time results [15], which enables an analysis of the surrounding environment, acting in accordance with the previously defined goals. AI agents play a pivotal role, since they demonstrate a great capacity and efficiency in data processing, on large scale [3]. In the future, AI is likely to considerably change both marketing strategies and customer behaviors [15]. AI assumes great importance for brand management, mapping the customers' journey. It subsequently originates from the data analysis, the marketing and sales conversion funnel, permitting to observe and drive the consumer, delivering the right information, services and promotions, in the right time and situation [3].

Although technologies are a generic lever for the creation of innovative strategies, the tourism sector presents a set of unique opportunities and challenges that should be considered when approaching and choosing technological strategies.

3 Tourism in Tech and Immersive Experience

In most sectors, the product becomes the experience, that is, it is important to have good products and good services, but above all, to provide a good experience, because the current and future generations are not interested in mass sales attempts made by the brands, but rather by living different experiences. Brands will have to address people by appealing to their emotions and creating unique sensations, instead of being based only on reason, since most consumption reasons are emotionally and subconsciously based [16].

Experiences constitute the essence of the tourism industry. Still, while the literature has recognized the recent impact of technology on experiences, its empirical exploration remains scarce [17]. Several examples of the applications of AR and VR techniques can be found in stores, tourism, hotel, restaurants, and destinations [18]. And immersive virtual environments, based on existing or upcoming technologies, are changing the way tourism operators stimulate their customers before, during, and after their experience [18, 19]. Yet, an extensive analysis of studies employing such techniques in tourism-related studies is difficult to find [18]. Loureiro et al. [18] provide and discuss the overall studies conducted between 1995 and 2019 on the application of VR and AR in tourism and suggest the most influential articles so far.

There seems to be potential for immersive technology, such as AR and VR, to attract tourists [20], to improve experiential outcomes [18, 21] and to create memorable tourism experiences, for heritage tourism [22], cruise tourism [18], thematic tourism/port wine [23] and many others. Moreover, rapid advancements in software and hardware, as well as significant investments in current technology, have made the application of AR and VR readily accessible [22]. VR and AR have undergone technical evolutions over the last few decades including improvements in immersion and the feeling of telepresence [18].

In tourism, Guttentag [24] refers that VR offers many practical applications that deserve greater attention from tourism researchers and professionals as VR technology continues to evolve, the number and significance of such applications undoubtedly will increase. Planning and management, marketing, entertainment, education, accessibility, and heritage preservation are six areas of tourism in which VR may prove particularly valuable.

Useful information in 3D tourism sites such as the visual resemblance and overall feel of the physical tourism destination, naturalistic elements, and images of cultural authenticity can enhance the consumer experience of enjoyment [20] and satisfaction with memorable experiences [18]. Three-dimensional virtual worlds provide opportunities for destination marketing organizations to communicate with targeted

markets and potential visitors, creating destination awareness in a rich environment [24].

The use of new technologies, although it is still scarce in tourism literature, is gaining momentum and will be the future [18].

In this context, the design of new models capable of maximizing the use of different technologies, according to a holistic view, appears as an opportunity to differentiate tourist destinations, making them not only technologically smart destinations but above all focused on providing smart information and services.

4 Proposed Conceptual Model

As mentioned in the previous chapters, there is no doubt that recent technology developments will increasingly enhance the promotion of territories and the ability to develop new and innovative marketing strategies. However, the existence of technology, in itself, is no guarantee of success. It is imperative that technology is fully interconnected and at the service of destinations/territories and their visitants.

The concept of smart cities emerges as a strategy to improve the quality of life of citizens and visitors [25], making use of a set of new technologies, including IoT, sensory devices, big data, wireless sensor networks, geospatial technology, among others [26].

In our view, an adequate approach is the vision of technological interconnection materialized in the so-called smart cities where this interconnection is enabled by IoT. As such, according to [27], smart cities are an important domain that is currently very much in vogue and where IoT plays an important role in its implementation. In this way, technology is, according to [28], what makes a city smart, in the view of many people. Also, studies carried out in the field of smart cities show that smart cities attract an increasing number of tourists [29].

Assuming that a smart city is a space in which technology is embedded in the overwhelming majority of its physical spaces, highly interconnected and at the service of several computational systems with multiple purposes for different players, Fig. 1 proposes a conceptual model capable of materializing the concept of immersive marketing, with a focus on tourism. The proposed model begins by identifying the three moments of a tourist visit—before, during, and after the visit. This is the cycle of a visit to a given destination.

Before a visit, a tourist is hungry for information about the destination he wants to choose or has already chosen. As such, it uses a wide range of information sources to satisfy its doubts/interests. It is here that tourists will naturally resort to social networks and/or specialized websites. In the proposed model, this is the first stadium where information about the profile of the potential visitor should begin. In this research process, the tourist will give a set of indications about their profile, which is materialized by their choices when searching, questioning, and interacting with various sources of information. It is at this stage that “the destination” should get to

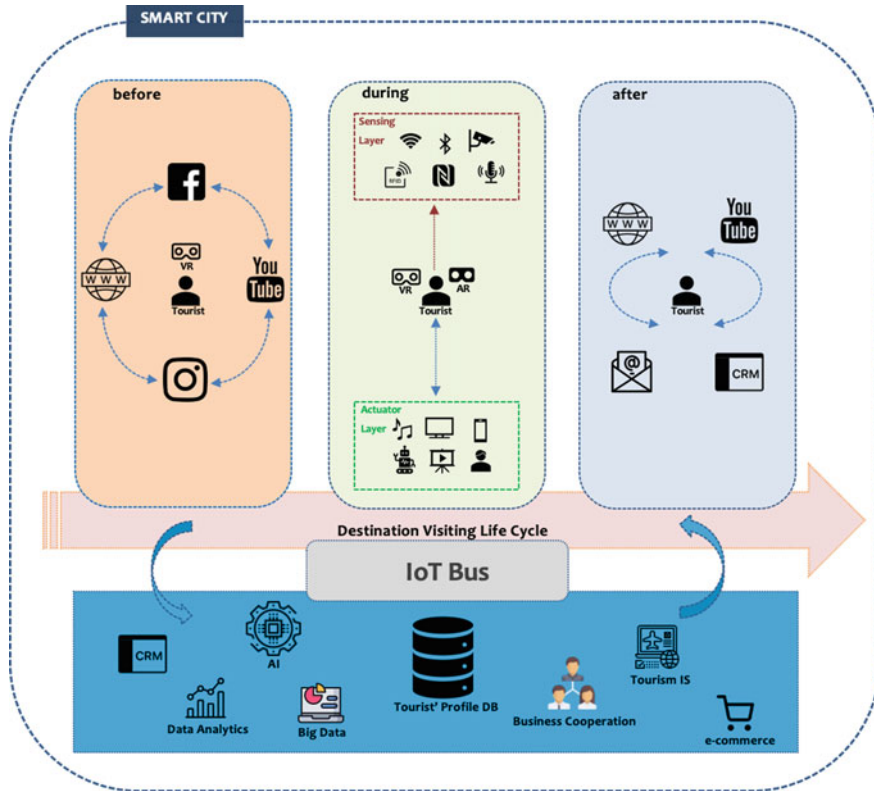


Fig. 1 Proposed conceptual model of an immersive marketing strategy for destinations

know the tourist and his profile. VR can play the upmost important role in this stage, innovating with an immersive experience of knowing a destination before visiting it.

During the visit, the visitor is hungry for information about everything he sees and will be increasingly demanding about the mechanisms at his disposal to interact with the spaces and get answers to his doubts or concerns—with a high level of personalization. A smart city (through all its embedded technology) must be able to use its sensory layer to identify and recognize the visitor trying to anticipate and match their desires (e.g., for information, services), using their layer of actuators (e.g., audio, video, Apps) to provide information and/or access to services in a personalized way and tailored to the visitor’s profile, circumstance, and location. It should be noted that this will only be possible because in the previous stadium (before the visit) information about the visitor’s profile was collected. In addition, all interactions that the tourist makes with the system should be used to improve their profile and to improve the intelligence quality of the system as a whole. It is at this stage that “the destination” begins to respond effectively and innovatively to a visitor who “in practice already knew” and that as there are interactions, get to know more and better. This stage generates satisfaction in the visitor who feels welcomed.

Finally, after the visit, the visitor wants to remember, share, and renew the experience of the visit. This phase is extremely important because it will tend to generate the opinion that the visitor had that will influence their future travel decisions, as well as everyone with whom the visitor shares their emotions. It is at this stage that “the destination” will be able to assess the degree of effectiveness and efficiency offered to the visitor—the degree of success of the visit.

It is important to mention that these three moments are stages of a visit cycle. And, as such, they feed in perpetual motion. Improving with each visit, with each visitor, with each common profile.

The basic component of the proposed model (IoT and all the technology that is connected to it) represents the support that allows the model to be able to assimilate information about the visitor, building his initial profile (before the visit); be able to process and combine visitor profile information and interactions in order to respond appropriately and personalized to the visitor (during the visit); as well as being able to use the information generated after the visit to understand the degree of success of the visit—the degree of success “of the destination”. All the main players in the destinations are associated with this technological combination (e.g., travel agencies, influencers, tourist companies, hotels, managers of cultural spaces). This layer is the materialization of a vast network of cooperation that, however, may itself be capable of generating business opportunities that will enhance the creation of new cooperation networks that will satisfy them.

5 Conclusion

The changes that have taken place in recent years, both in society and in business, have moved towards a technological revolution, in general, and particularly, in organizations, especially in the way they manage their relationship with their target audience. This reality has led companies to become more agile, virtual, and collaborative.

In the near future, in addition to a digital revolution, entities and the market will also undergo a cultural revolution, and the relationship with customers must be increasingly human. Words such as Onlife (merge online and offline), One-to-Moment, Human-to-Human, Emotion Marketing, and Human Touch will be present in marketing strategies, as they allow empathy to build a close relationship between brands and customers [16].

This reality becomes challenging for multiple areas of knowledge—marketing is no exception.

This technological revolution, where different technologies are increasingly interconnected and embedded in physical spaces, has allowed the creation of smart spaces, its greatest example being the creation of smart cities. In this context, Marketing—which, par excellence, lives at the expense of maximizing the interaction of the trinomial company/product/customer, will be able to extract minimal potential from the aforementioned technological revolution. There was never, like today, such a great capacity to follow and understand the aforementioned trinomial.

In the context of tourism and tourism marketing, new opportunities open up to support the life cycle of a visit to a given destination in an innovative and effective way. Knowing the profile of the potential visitor (before), responding effectively, intimately, and assertively during the visit, and being able to measure the success of the visit (after), reveals the enormous potential that marketing has today to make a CRM suitable for promoting destinations and generating business opportunities.

This article, after framing the context of digital marketing under the scenario of the technological revolution of recent years and under the aegis of smart cities, reflects on the potential of recent technologies in the context of tourism and the cycle of a visit to a given destination. Finally, it presents a conceptual model that presents itself as a manifesto for reengineering of thought that, today, can be operationalized and that would greatly help in the sustainability and business creation of destinations and their territories, as well as in the degree of satisfaction of the visitors. However, the proposed model requires a dense network of collaboration so that it can be operationalized—here lies an important limitation of the work presented.

Acknowledgements UNIAG, R&D unit funded by the FCT—Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education. UIDB/04752/2020.

References

1. Cluley, R., Green, W., Owen, R.: The changing role of the marketing researcher in the age of digital technology: practitioner perspectives on the digitization of marketing research. *Int. J. Mark. Res.* **62**(1), 27–42 (2020)
2. Kotler, P., Kartajaya, H., Setiawan, I.: *Marketing 5.0: Technology for Humanity*. Wiley (2021)
3. Ma, L., Sun, B.: Machine learning and AI in marketing—connecting computing power to human insights. *Int. J. Res. Mark.* **37**(3), 481–504 (2020)
4. Hudson, S., Matson-Barkat, S., Pallamin, N., Jegou, G.: With or without you? Interaction and immersion in a virtual reality experience. *J. Bus. Res.*, **100**, 459–468 (2019)
5. Navío-Marco, J., Ruiz-Gómez, L.M., Sevilla-Sevilla, C.: Progress in information technology and tourism management: 30 years on and 20 years after the internet—revisiting Buhalis & Law’s landmark study about eTourism. *Tour. Manag.* **69**, 460–470 (2018)
6. Bethapudi, A.: The role of Ict in tourism Industry. *J. Appl. Econ. Bus.* **1**(4), 67–79 (2013)
7. Tian, F.: Immersive 5G virtual reality visualization display system based on Big-Data digital city technology. *Math. Probl. Eng.* **2021**, 6627631 (2021)
8. Leopardi, A., et al.: X-reality technologies for museums: a comparative evaluation based on presence and visitors experience through user studies. *J. Cult. Herit.* **47**, 188–198 (2021)
9. OECD: *OECD Tourism Trends and Policies 2020* (2020)
10. Syed, A.S., Sierra-Sosa, D., Kumar, A., Elmaghraby, A.: IoT in smart cities: a survey of technologies, practices and challenges. *Smart Cities* **4**(2), 429–475 (2021)
11. Jasrotia, A., Gangotia, A.: Smart cities to smart tourism destinations: a review paper ... *Tour. Intell. Smartness*, pp. 47–56 (2018)
12. Kumar, V., Ramachandran, D., Kumar, B.: Influence of new-age technologies on marketing: a research agenda. *J. Bus. Res.* **125**, 864–877 (2021)
13. Lai, Z.: Research on advertising core business reformation driven by artificial intelligence. *J. Phys. Conf. Ser.* **1757**(1), 012018, (2021)

14. Sözer, E.G., Civelek, M.E., Çemberci, M.: Strategic Excellence in Post-digital Ecosystems: A B2C Perspective. E-Books, Zea (2018)
15. Davenport, T., Guha, A., Grewal, D., Bressgott, T.: How artificial intelligence will change the future of marketing. *J. Acad. Mark. Sci.* **48**(1), 24–42 (2020)
16. Oliveira, C.M.: Marketing Pós-Digital: O marketing à medida do ser humano. Actual Editora (2020)
17. Neuhofer, B., Buhalis, D., Ladkin, A.: A typology of technology-enhanced tourism experiences. *Int. J. Tour. Res.* **16**(4), 340–350 (2014)
18. Loureiro, S.M.C., Guerreiro, J., Ali, F.: 20 years of research on virtual reality and augmented reality in tourism context: a text-mining approach. *Tour Manag.* **77**, 104028 (2020)
19. Lemon, K.N., Verhoef, P.C.: Understanding customer experience throughout the customer journey. *J. Mark.* **80**(6), 69–96 (2016)
20. Huang, Y.C., Backman, K.F., Backman, S.J., Chang, L.L.: Exploring the implications of virtual reality technology in tourism marketing: an integrated research framework. *Int. J. Tour. Res.*, **18**, 116–128 (2016)
21. Wu, H.C., Ai, C.H., Cheng, C.C.: Virtual reality experiences, attachment and experiential outcomes in tourism. *Tour. Rev.* **75**(3), 481–495 (2020)
22. Bec, A., Moyle, B., Timms, K., Schaffer, V., Skavronskaya, L., Little, C.: Management of immersive heritage tourism experiences: a conceptual model. *Tour. Manag.* **72**, 117–120 (2019)
23. Martins, J., Gonçalves, R., Branco, F., Barbosa, L., Melo, M., Bessa, M.: A multisensory virtual experience model for thematic tourism: a Port wine tourism application proposal. *J. Destin. Mark. Manag.* **6**(2), 103–109 (2017)
24. Guttentag, D.A.: Virtual reality: applications and implications for tourism. *Tour. Manag.* **31**(5), 637–651 (2010)
25. Osman, A.M.S.: A novel big data analytics framework for smart cities. *Futur. Gener. Comput. Syst.* **91**, 620–633 (2019)
26. Ahad, M.A., Paiva, S., Tripathi, G., Feroz, N.: Enabling technologies and sustainable smart cities. *Sustain. Cities Soc.* **61**, 102301 (2020)
27. Janani, R.P., Renuka, K., Aruna A., Lakshmi Narayanan, K.: IoT in smart cities: a contemporary survey. *Glob. Trans. Proc.* (2021)
28. Zhao, F., Fashola, O.I., Olarewaju, T.I., Onwumere, I.: Smart city research: a holistic and state-of-the-art literature review. *Cities*, **119**, 103406 (2021)
29. Eichelberger, S., Peters, M., Pikkemaat, B., Chan, C.-S.: Entrepreneurial ecosystems in smart cities for tourism development: from stakeholder perceptions to regional tourism policy implications. *J. Hosp. Tour. Manag.* **45**, 319–329 (2020)

Digital Marketing and Branding

Are Colors Emotional Triggers in Digital Branding?



Joana Sampaio Correia and Dora Simões

Abstract This investigation exhibits a real study of the influence of color in a digital promotional product—social coupon. The paper’s goal is to analyze the perception of color in graphic design, through the analysis of consumers’ responses. Following the design-science research (DSR) approach through the development of the artifacts—social coupons—the emotional influence of color was studied during the literature review. Then, according to the DSR, the main problem was divided into iterative cycles, in order to build and evaluate the artifacts designed to meet the identified business need. After the development of innovative artifacts, a digital promotional product was tested and redesigned, according to the user’s perceptions. Consumers prefer colors that are associated with positive emotions and find tranquility and safety the most related to blue color, sophistication to black color, comfort to white color, brown as the most boring color, and red as the color which captures more of their attention. This study brings new insights to brands, with the development of a new framework and with its exploration about the emotional triggers of colors in the design, determining how to best capture consumers’ attention and how to bring more pleasant experiences to them.

Keywords Color psychology · Digital branding · Aesthetics · Emotional design · Digital promotional product · Social engagement

1 Introduction

Digital technological mediation in the purchasing market allowed for greater competition, leading brands to adopt new solutions and strategies that would lead them to stand out in the public [17]. As part of the new digital strategies to acquire customers,

J. S. Correia (✉)

Department of Communication and Art, University of Aveiro, Aveiro, Portugal

e-mail: joanamscorreia@ua.pt

D. Simões

Digital Media and Interaction Research Centre, Higher Institute of Accounting and Administration, University of Aveiro, Aveiro, Portugal

e-mail: dora.simoes@ua.pt

social coupons emerged. Social coupons, as the name implies, are online discount vouchers, which are available from a variety of companies through websites of daily offers. Also labeled as “flash sales” or “daily deals”, social couponing is a form of sales promotion that has enjoyed widespread popularity among hotels in the last few years [14, 6]. Different from traditional coupons, social coupons not only have the potential to promote short-term sales but maybe also used to increase brand identity and awareness [6].

In business, interaction with consumers is typically understood as a process, not a long-term relationship. By managing communication in both directions properly, a relationship can be long-term. This tactic determinates the creation of a strong relationship between consumers and brands with a two-way dialogue, communication, and cooperative interaction [4]. Thus, social engagement is a positive motivational state in work and is characterized by passion, dedication, and immersion [9]. Zheng et al. [28], reinforce that engaged people have confidence and support the organization’s goals, have a sense of ownership, feel proud of the organization, and desire to develop and survive in the organization. Therefore, the company must tactically make the organization survive and improve. Marketers need to understand engagement through the company/consumer relationship.

According to Gobé [12], sensory experiences are powerful, immediate, and capable of changing people’s lives. So, product development must be developed to awaken different senses in the consumer, stimulating at a visual level, by using something aesthetically pleasing and intuitive, mixing differentiated and bold textures and shapes, making the product irresistible to everybody’s eyes. Thereby, emotional design is essential to capture the consumer’s attention.

Following the design science research (DSR) methodology, this paper is structured as follows:

1. Introduction.
2. The aesthetics of digital branding.
3. Design science research.
4. Artifacts—social coupons proposal.
5. Evaluation and artifacts re-design.
6. Final remarks and business implications.

2 The Aesthetics of Digital Branding

According to Labrecque and Milne [18], color attracts consumers and may shape their perceptions, acting as a marketing tool. Through color, a brand may be able to establish a visual identity, create strong relationships with its target audience and position itself vis-à-vis its competitors in the market.

2.1 The Emotional Side of Involvement

Vaughn and Leff [25], investigated the emotional component of the involvement perceived in the product categories, outlining a theoretical perspective to envision it. Therefore, they added a second orthogonal dimension to the concept of high and low involvement. The second dimension added is “thinking versus feeling”. This framework (Fig. 1) is based on the traditional model, which implies that high involvement products require an early cognitive orientation, while low involvement products are more suitable for an affective or non-informative appeal [10]. Expanding engagement over an orthogonal continuum of thought allows for a more complex approach, in which perhaps the excitement that accompanies certain purchases may be taken into account [20].

The first quadrant corresponds to the greatest involvement/thinking and implies a great need for information due to the importance of the product and to the issues of thinking about it. The second quadrant involves the decision of the product, but the specific information is less important than a holistic attitude or feeling towards the product. The third quadrant corresponds to low involvement/thinking and decisions about products in this area, which may require minimal thinking and a tendency to form buying habits for convenience. The fourth quadrant is low affective involvement and is reserved for products that satisfy personal tastes [25].

The challenge for advertisers is that each quadrant has its peculiarity. Therefore, they require different types of advertising. This concept is intuitively attractive because it allows the emotional aspects often associated with certain products to be recognized and incorporated into the advertising strategy [20]. Different products require different needs. Consequently, they require new marketing strategies

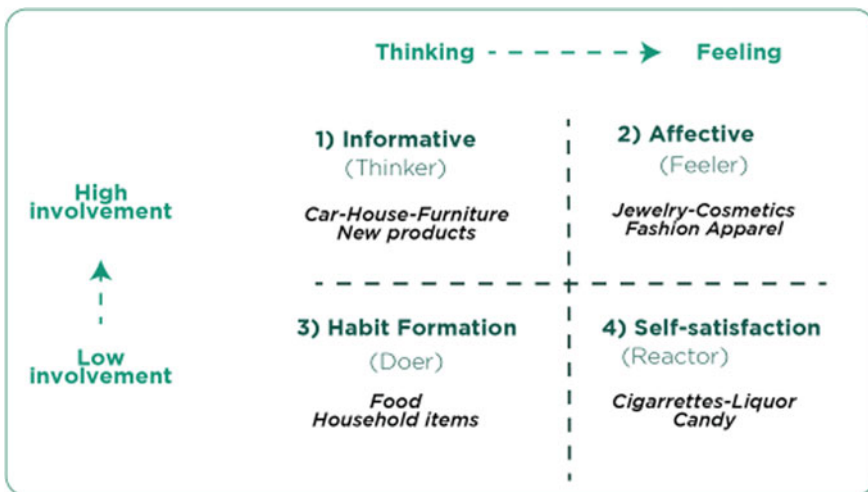


Fig. 1 Framework of the emotional side of involvement. Adapted from Vaughn and Leff [25]

focused on visual stimulus. Thereby, aesthetics become a new marketing model that guarantees companies the possibility of approaching consumers through a wide variety of sensory experiences. Thereupon, it benefits both consumers and brands by promoting customer satisfaction and loyalty [8]. The design and aesthetic quality of the product are increasingly important in the market, where many of the basic needs of customers have already been known. As the essential attributes of the products (such as quality and functionality) become more and more homogeneous, companies resort to products' less tangible resources, so that they become recognizable [22].

Nowadays, aesthetic is the element that makes the distinction between products, according to the fact that technical functions are easily copied [24]. It is essential to create new products based on their visual qualities and their appealing aesthetics. The best way to accomplish success in companies is closely related to the concept of aesthetics because, in addition to influencing consumer behavior, it will also affect corporate identity and brand image [21]. For example, the ZARA Company (Fig. 2), one of the main brands of the INDITEX Group is a successful model that managed to use aesthetics from the beginning, as an advantage to create its own identity. As such, through the company's name, logo, website and own style, corporate and brand expressions, it is possible to illustrate how the company managed to create a positive image among consumers, using aesthetics [8]. In March 2020, due to the Coronavirus lockdown, ZARA Company discovered a visual strategy adapted to this phase (Fig. 2), through a short animation, using the increase and decrease of the tracking as a metaphor for social distance, creatively and positively.

It should also be noted that the visual aspect influences the perception and consequent emotion in those who spot the product for the first time. As such, many brands try to provoke reactions and feelings in the consumer through an attractive design that stimulates triggers in those who are perceiving the product [15]. Color, shape, packaging, design, and aesthetics will arouse emotions and consequently bring "desire" and attraction for the product. A good example of the aesthetic influence on consumer choice, more specifically in pop art, is the well-known cans of tomato soup from Campbell's. Through the repetition and uniformity of advertising, Andy Warhol reproduced in detail the same image on each screen, imitating the characteristic repetition of mass production and the uniformity of advertising, modifying only the label related to the variety of each can [15]. Thanks to Andy Warhol, Campbell's iconic design has become a reference in pop art forever.

Similarly, to Campbell's Soup Cans, Coca-Cola bottles are also a great example of how a visually interesting aesthetic can captivate and win back the public. It is therefore also considered an icon of mass culture. The Coca-Cola bottle's exclusive concept emerged from the defensive marketing strategy carried out in 1915 when

Fig. 2 Animation of the ZARA Company, during the Covid outbreak (<https://www.zara.com>)



the brand was losing market share and therefore decided to launch a national competition for a new bottle design. Its bottles were strategically designed, so, in case of breakage, its pieces would allow it to recognize that it was a bottle of Coca-Cola and its shape should be so particular that it could be easily recognized in the dark [1]. Given the exclusivity of the bottle, consumers would realize that Coca-Cola was a premium product that should not be confused with any other coke presented in a transparent glass bottle. This characteristic and original packaging represent the brand in an aesthetically simple, minimalist, and classic way. That's why the consumer is attracted to the brand [23].

Beardsley [3], argued that experiences and aesthetic values are linked to emotional reactions. Therefore, he emphasized that it is necessary to make an in-depth study of emotions. Starting from the need to explore emotions, the terms “neurodesign”, “neurobranding” and “neuromarketing” have emerged. According to Krajnovic et al. [15], neurodesign may be combined with neurobranding, focusing on the study of emotions, trying to awaken desired emotions in consumers through colors, scents, design, and music. It is essential to create a synergy of all sensations. A unique experience where all the senses are represented, and which so strongly affects the decisions of everyone. By stimulating all the senses, you will start a sequence of unstoppable associations, creating a new emotional state. In the authors' opinion, emotion is the new “weapon” of the twenty-first century.

2.2 *Color Psychology*

Color is an important variable that affects the consumer's perception in advertising, creating associations and visual stimulus. Table 1 summarizes the influence of color on the consumer's mind, through the colors: white, yellow, orange, red, pink, purple, blue, green, brown, grey, and black. The first column identifies the meaning associated with each of the colors mentioned. In the second column, tones are characterized according to their warmth: warm, neutral, or cold [7].

Thus, the colors yellow, orange, red, and pink are identified as warm tones; the colors purple, blue, and green as cold tones and finally, the colors white, brown, grey, and black are identified as neutral tones. Then, the third column shows which colors are associated with positive or negative emotions. Warm tones lead to positive emotions and cold tones to negative emotions. Regarding neutral tones, the emotions associated with them may vary, according to the intensity of the color. For example, if the color is light grey, it can transmit neutral to positive emotions, while darker grey and black colors will bring negative characteristics [7]. In the next column, the colors are classified according to their intensity, exploring if a certain color can confer a high, moderate, or low significance [27]. Black and white are considered pure solid colors, thus, they are not evaluated through this category. The last column presents some affiliations to certain colors, resulting from individual and cultural experiences [13]. Examples of these cultural experiences presented are the association of white color with purity and religion, while black is associated with death and funerals. On

Table 1 The emotional influence of color on consumers

Color	Meaning	Temperature	Emotion	Intensity	Association (individual and cultural experiences)	Brands
White	Serenity Peace Simplicity	Neutral	Positive	————	- Purity - innocence - Angels of peace - Religion	
Yellow	Happiness Sincerity Optimism Extroversion	Warm	Positive	Medium	- Luck	McDonald's
Orange	Sunny Energy Relaxation Socialization	Warm	Positive	High		Pouch
Red	Love Madness Passion Enthusiasm Strength	Warm	Positive	High	- Hitchcock horror movies, - S. Valentine's marketing campaigns, - Coca-cola	Continente Beruby
Pink	Feminism Calm Sweetness Sincerity Tenderness	Warm	Positive	Low	- Color associated with the feminine by traditional culture	
Purple	Sophistication Mystical Luxury Authenticity Quality	Cold	Negative	High	- Color reserved for royalty to connote social roles in the past	Retail Me Not
Blue	Comfort Safety Competency Reliability Duty	Cold	Neutral/ Negative	Negative feelings are increased with color intensity	- Color associated with masculine by traditional culture - Pepsi brand.	VoucherCloud
Green	Nature Calm Safety	Cold	Positive	Medium	- St. Patrick's Day Lucky Clovers (Ireland)	Compal
Brown	Nature Earth Respect Trustful	Neutral	Neutral/ Negative	Low		
Grey	Boring Monotonous	Neutral/ Absence of emotion	Neutral/ Negative	Low		
Black	Sinister Modernism Fashion Elegance Glamour	Neutral	Negative	————	- Association with death and funerals, - Color of limousines, tuxedos, black ties at events	PROZIS

Adapted from Labrecque and Milne [18], Clarke and Costall [7], Grossman and Wisenbilt [13], and Wexner [27]

the other hand, there are colors culturally associated with luck, such as yellow, due to the shooting stars and the wish requests associated with this astronomical event. In addition, green can also be related to luck, according to the lucky clovers and the Irish tradition of St. Patrick's Day. Other colors are related to terror and blood, as is the case of the association between the color red and Hitchcock's horror films. On the other hand, red can also relate to love and Valentine's Day marketing. It is also relevant to mention which colors are associated with the gender: pink for the female and blue for the male, according to the traditional culture [18]. Finally, the last column specifies brands' colors of social coupons and their association with each of the colors under study.

3 Design Science Research

The implementation of design science research (DSR) needs to be both objective and systematic [26], as with all investigations. According to this thought, DSR is parallel to instructional design in several ways. This methodology aims to answer questions related to the usefulness of the developed artifacts and their motivations (“What use does the new artifact provide?” and “What demonstrates this usefulness?”). The core of the design sciences is its focus on utility [11, 5].

3.1 *Development of an Innovative Purposeful Artifact*

As stated earlier, the design of an artifact, its formal specification, and an assessment of its suitability, often in comparison to competing artifacts, are essential for the investigation of design science. These should be combined with behavioral and organizational theories, to develop a deep understanding of the problems, contexts, solutions, and approaches to business assessment, to serve the research communities [11]. This methodology is considered a theory of design, so, it is essential to detail the relevance of its cycles. The design cycle begins with the awareness of the problem or situation, which, as a rule, is usually large and complex [16]. Then, the model is evaluated and refined in consequent design cycles. The evaluation method of these cycles is iterative, so the evaluation is not considered the final stage: the circumscription/limitation that occurs after development and the consequent evaluation is an extremely valuable stage [5]. These design sciences represent real-world problems and are focused on the creation and evaluation of new products/artifacts. Thus, it has its core on their design, implementation, and evaluation, and on the understanding of previously created artifacts, according to their experimental and innovative design. This methodology is inclusive and iterative, aimed at the development of digital artifacts, which are objects created by humans and part of the sciences of the “artificial” [19].

Hereupon, the first phase of the research was focused on the awareness of the problem and the analysis of the main aspects related to design and their influences on consumer’s preferences, as presented in Fig. 3.

A questionnaire was developed, to obtain data about consumer color preferences. This questionnaire had 152 participants from all digital generations: 69.7% female and 30.3% male; 53.3% from Gen Z (1995–2010); 20.4% from Gen Y (1982–1994); 19.1% from Gen X (1965–1981); 7.2% from Baby Boomers (1946–1964). The questionnaire was divided into four different sections. The first one was focused on the emotional influence of color on consumers, through the literature review performed. Thus, 9 color palettes of different brands that use social coupons were presented: Retail Me Not, VoucherCloud, Coupons.com, Pouch, Beruby, Prozis, Continente (Sonae), and McDonald’s. The second section of the questionnaire displayed two colors: red and black and asked the inquired to select which of the presented emotions

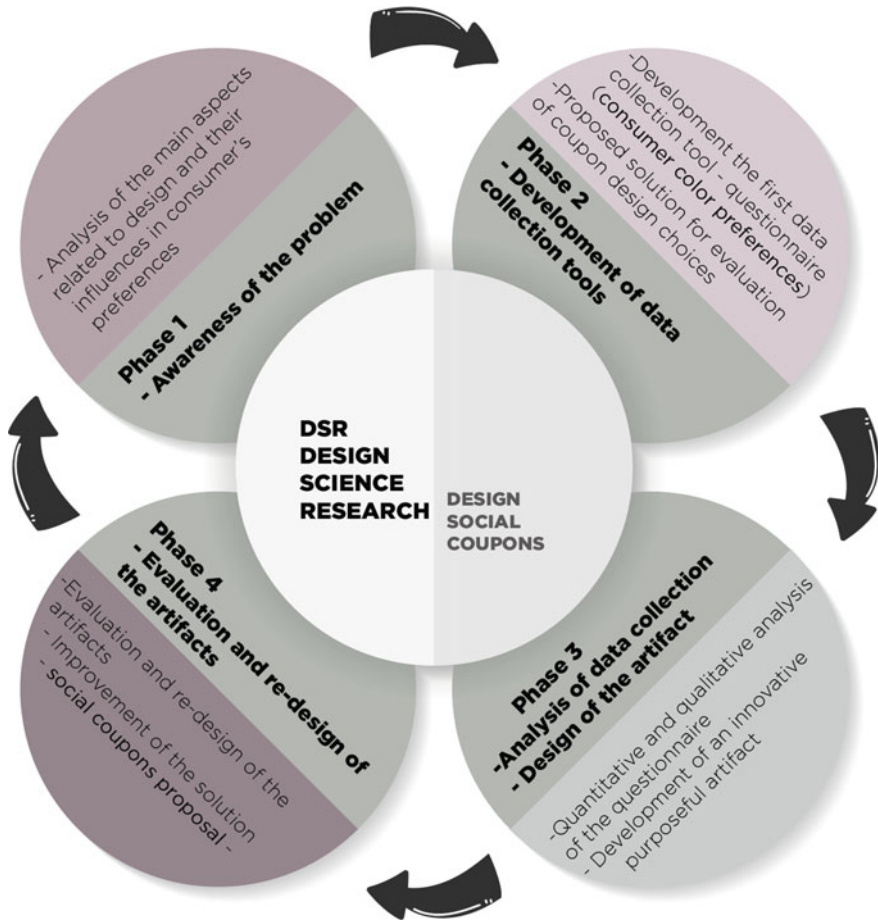


Fig. 3 Design science research

was the most related with each color, exploring the duality of the two colors. The third section explored the associations between colors and brands, through their emotional design. Thus, the inquired had to write which brand they associated with each color: red and yellow. The fourth and last section presented 6 emotions: tranquility, sophistication, comfort, safety, boredom, and attention, and it was asked which color was associated with the listed emotions.

Data were analyzed in SPSS (24 Version) and in Microsoft Excel (16.39 Version). Afterward, the artifacts were designed with Adobe Illustrator 2019.






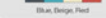
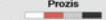

3.2 Evaluation and Re-design of the Artifacts

The artifacts were evaluated through a first questionnaire directed to consumers, as mentioned previously. This questionnaire studied consumers preferences’ through four different groups of questions: color, the duality of color, the color of brands, and emotional design. A group of simple coupons—social coupons—with different colors and shapes was also presented. The emotional side of design was explored through a social engagement tool: social coupons. First, respondents chose the most attractive color palettes on the first question. Then, they decided if red and black were most associated with positive or negative emotions. On the third question, they responded spontaneously, mentioning which brand was more related to the colors: red and yellow. On the last question of the questionnaire, the respondents had to link the emotions: tranquility, sophistication, comfort, safety, boredom, and attention with the colors listed: white, pink, blue, green, brown, grey, red, purple, black, yellow and orange.

4 Artifact Description—Social Coupons Proposal

The artifact was developed, according to the framework presented in the last section (see Table 2): a social coupons’ proposal. The grey shadows show consumers’ answers. According to the color palettes presented, consumers have a strong preference for the color palettes of the brands: coupons.com, Beruby, and Prozis. The three-color palettes have warm tones. The first one has lighter tones of orange and yellow, whereas both Beruby and Prozis have a darker tone of red. The color palettes

Table 2 Framework of the results

Color (Literature review)	Duality of color (Colors with contradictory interpretations)	Color of brands (Associations between brands and colors)	Emotional design (Color psychology)
RetailMeNot  Purple, Light Orange, Red	Red - Anger - Passion	Red - (Continente (Sonae)) - (Globo (Cob)) - Ferrari - McDonald's - Prozis - Benefica - Lewis - Netflix - Worten - EDP - K&T Kat - Vodafone - CTT Yellow - McDonald's - Shell - Leif - Fnac - DR.L - BEA - Benetton - CAT - Ferrari - Lays - Lipton - Nesquik - Opel	Tranquility - White - Pink - Blue - Green - Brown - Grey Sophistication - Red - Pink - Purple - Grey - Black Comfort - White - Yellow - Pink - Blue - Green - Brown - Grey - Black Safety - White - Black - Green - Brown - Grey Boredom - White - Black - Yellow - Grey - Black Attention - White - Yellow - Orange - Purple - Pink - Blue - Green - Brown - Grey - Black
VoucherCloud  Blue, Orange, White	Black - Targeted/Grander - Fashion/Glamour		
coupons.com  Green, Blue, Yellow, Orange, Grey			
Pouch  Orange, Green, Blue, Grey, White			
Beruby  Black, Orange, Red			
Prozis  White, Red, Grey (light and dark)			
Continente  White, Red, Grey, Black			
Mc Donald's  Yellow, Red, Green, Brown Pink, Black, Green, Orange			

of Coupons.com and Beruby both have the color blue, while Prozis is more centered on the color red, complemented with black, white, and grey. Concerning the duality of the color section, consumers chose positive emotions as much to red color, as for the black color. The results of the third section show that the color red is particularly associated with Continente and Coca-Cola, whilst yellow color is substantially referred to McDonald’s. The last section shows that consumers identify the blue color to tranquility; black to sophistication; white to comfort; blue to safety; brown to boredom and red as the one that captures their attention.

According to the fact that respondents chose color palettes of the brands “coupons.com”, “Beruby” and “Prozis”, they showed a predisposition to warm tones and light shades of blue and green, with grey gradients. Thus, it would be relevant to bet on the design of coupons where these colors were used.

The creative process of the artifacts started with some sketches. First, a simple and flat-shaped coupon, with different corners—straight or rounded—was drawn. Then, other similar coupons with different outlines were designed. After that, color and shadows were added to the coupon, for it to look real. Based on the results acquired during data collection (Table 3), a proposal of 9 coupons was developed with different colors, illustrations, and lettering. Similarly, to traditional coupons, the final version of the proposal is horizontal. The proposal is presented in Fig. 4.

The coupons numbered 4, 5, and 8 were drawn with warm tones of red, orange, and yellow, with simple and significant illustrations, related to the discount theme. Coupon number 2 shows a gradient between warm and cold tones, softening the color blue to orange. Then, two other coupons were developed: 6 and 9. These coupons were designed using “comfortable” and “safe” colors: blue and green. Also, they have a little texture (coupon number 6—“Mattress”) and some abstract illustrations (coupon number 9—“Seal”). Then, another 3 coupons were developed: 1, 7, and 3. These coupons mostly have warm tones, funny and soothing illustrations, with some shades and experimental themes.

Table 3 Evaluation of the artifacts

Digital Age - Segmentation												
Focus Group results												
	Baby Boomers (1946-1964) [66-74]			Gen X (1966-1981) [89-95]			Gen Y (1982-1994) [26-38]			Gen Z (1995-2010) [10-25]		
	♀ ♂		♀ ♂		♀ ♂		♀ ♂		♀ ♂		♀ ♂	
	2	1	1	1	1	1	1	1	1	1	1	1
1st	2	1	1	1	4	5	9	5	9			
2nd	4	4	6	4	1	6	4	4	4	4		
3rd	7	5	7	7	6	9	6	2	7			

1 - Coupon "Mineral water" - 4x	4 - Coupon "Orange" - 7x	7 - Coupon "Tuna" - 4x
2 - Coupon "Simple" - 2x	5 - Coupon "Ongami" - 3x	8 - Coupon "Literature" ---
3 - Coupon "Shoe"---	6 - Coupon "Mattress" - 4x	9 - Coupon "Seal" - 3x

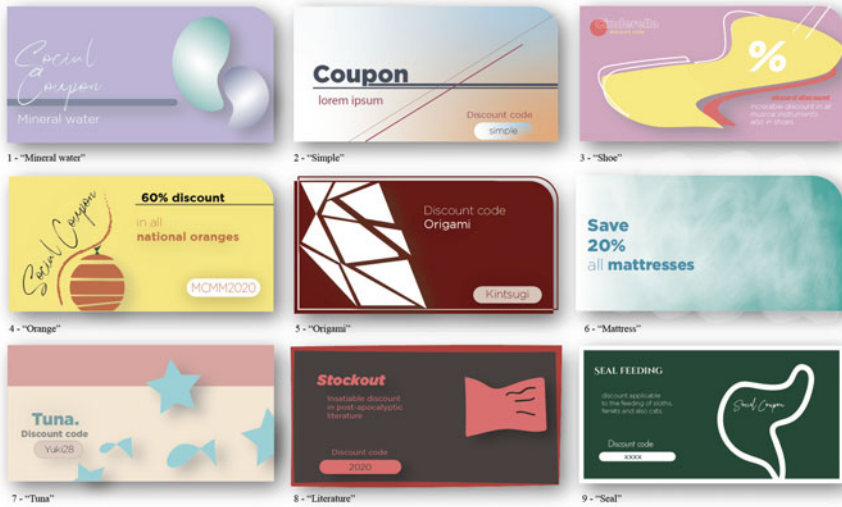


Fig. 4 Social coupons proposal

5 Evaluation and Artifact Re-design

Coupon proposals were evaluated through a focus group session validating them with 9 individuals of different generations and different genders (see Table 3).

The validation of the coupon proposals in the focus group session provided new insights to the study, showing the opinions and preferences of the participants. The coupons were presented in a different sequence, to avoid bias in the results.

In general, it appears that, regardless of gender and age, the coupon that most attracted everyone’s attention was coupon number 4—“Orange”. This coupon was mentioned 7 times, and only 1 individual classified it as 1st choice. This participant was male and generation X (see Table 3). On the other hand, this coupon was also classified 6 times as the 2nd choice, by individuals of different ages and genders, which shows a huge preference for this coupon in all generations. Participants’ preference for coupon 4 was justified by their color, illustration, content readability, and creativity. A female per-son, belonging to generation Y described the colors used as "cheerful and attractive, which capture our attention" and referred that "the colors match each other and motivate us to buy that product". Another individual, belonging to the female gender, and to the Baby Boomer generation, mentioned that the drawing itself attracted his attention the most: “very appealing design”.

In the end, new colored and vertical versions of the coupons previously designed were created based on the previously obtained feedback from the focus group (Fig. 5).

Thus, only five vertical versions were created, according to the top choices of the respondents during the focus group. These new versions are more adapted to digital platforms and may become responsive in mobile versions.



Fig. 5 The new social coupon's proposal

6 Final Remarks and Business Implications

This study led to new insights to brands and businesses that could be published through their social media, such as how to best capture consumers' attention and how to best develop digital promotional products for them. Also, the development of a framework shows best practices, and presents the relevance of emotion in design, in a digital age, whereas visual information is overloaded.

It would be interesting to make tests with a bigger number of consumers/users in the future and to assemble neuromarketing tests, to better understand consumers' expectations and motivations. Also, it would be relevant if brands had considered customers' feedback, privileging their comfort and a more intuitive and comfortable user experience, mixing traditional and digital strategies, displaying positive emotions through a simple yet beautiful design.

The social engagement tools—social coupons—showed a great method to evaluate the emotional side of design. The “desired” discount is like a “trigger” and may capture the customer's attention. This, its visual aspect, as the study shows, reflects that consumers are more engaged with colors that bring positive emotions, like red, orange, and yellow colors, and with safe colors like blue and green. Color psychology helps understand social engagement on the web. Colors are linked to feelings and decision-making, thus it is essential to create new products based on their visual qualities and their appealing aesthetics. Through an attractive design with different shapes and colors, a brand may incite reactions and feelings, stimulating triggers in those who are perceiving the product.

According to the main question: “Are colors emotional triggers in digital branding?”, it could be said that colors are emotional and engagement triggers in digital branding and may head consumers to their tangible needs, determining how to best capture consumers' attention and how to bring more pleasant experiences to them.

Acknowledgements Further research studies on this issue are detailed on a master's thesis entitled “The influence of design in the usage of social coupons” of University of Aveiro.

References

1. 11 Facts About the Coca-Cola Bottle: Retrieved November 21, 2019, from <https://www.coca-colacompany.com/au/news/11-facts-about-the-coca-cola-bottle.html> (2019)
2. Algharabat, R., Rana, N.P., Alalwan, A.A., Baabdullah, A., Gupta, A.: Investigating the antecedents of customer brand engagement and consumer-based brand equity in social media. *J. Retail. Consum. Serv.* **53**, 0–30 (2020). <https://doi.org/10.1016/j.jretconser.2019.01.016>
3. Beardsley, M.C.: Aesthetic experience regained. *J. Aesthet. Art Critic.* (1969). <https://doi.org/10.2307/428903>
4. Camilleri, M.: Market Segmentation, Targeting and Positioning. https://doi.org/10.1007/978-3-319-49849-2_4 (2017)
5. Carstensen, A.K., Bernhard, J.: Design science research—a powerful tool for improving methods in engineering education research. *Eur. J. Eng. Educ.* **44**(1–2), 85–102 (2019). <https://doi.org/10.1080/03043797.2018.1498459>
6. Cassia, F., Magno, F., Ugolini, M.: The perceived effectiveness of social couponing campaigns for hotels in Italy. *Int. J. Contemp. Hosp. Manag.* **27**(7) (2015). <https://doi.org/10.1108/IJCHM-02-2014-0090>
7. Clarke, T., Costall, A.: The emotional connotations of color: a qualitative investigation. *Color. Res. Appl.* **33**(5), 406–410 (2008). <https://doi.org/10.1002/col.20435>
8. Comiati, R.: The role of the aesthetics in the making of a brand. *Manag. Market.* **8**(2), 411–416 (2010)
9. Crittenden, V.L., Crittenden, W.F.: Digital and social media marketing in business education: implications for student engagement. *J. Mark. Educ.* **37**(3), 131–132 (2015). <https://doi.org/10.1177/0273475315588112>
10. Engel, J.F., Blackwell, R.D.: *Consumer Behavior*. Dryden Press (1982)
11. Esearch, S.Y.R., Hevner, B.A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. *MIS Q.* **28**(1), 75–105 (2004)
12. Gobé, M.: *Emotional Branding: The New Paradigm for Connecting Brands to People*. Windsor (2001).
13. Grossman, R., Wisenbilt, J.: What we know about consumers' color choices. *J. Market. Pract.: Appl. Market. Sci.* **5**(3), 78–88 (1999). <https://doi.org/10.1108/02652320210446724>
14. Katerina Berezina, K.J.S., et al.: The managerial flash sales dash: is there advantage or disadvantage at the finish line? *Int. J. Hosp. Manag.* (2016). <https://doi.org/10.1016/j.ijhm.2016.01.003>
15. Krajnovic, A., Sikiric, D., Jasic, D.: Neuromarketing and Customers' Free Will. *Managing Transformation with Creativity*. In: *Proceedings of the 13th Management International Conference*, pp. 1143–1163 (2012)
16. Kuechler, B., Vaishnavi, V.: On theory development in design science research: anatomy of a research project. *Eur. J. Inf. Syst.* **17** (2008). <https://doi.org/10.1057/ejis.2008.40>
17. Kumar, V., Rajan, B.: Social coupons as a marketing strategy: a multifaceted perspective. *J. Acad. Mark. Sci.* **40**(1), 120–136 (2012). <https://doi.org/10.1007/s11747-011-0283-0>
18. Labrecque, L.I., Milne, G.R.: Exciting red and competent blue: the importance of color in marketing. *J. Acad. Mark. Sci.* **40**(5), 711–727 (2012). <https://doi.org/10.1007/s11747-010-0245-y>
19. Lacerda, D.P., Dresch, A., Antunes Júnior, J.A.V., Proença, A.: Design science research: a research method to production engineering. *Gestao e Producao* **20**(4), 741–761 (2013). <https://doi.org/10.1590/S0104-530X2013005000014>
20. Lynne, J.: The emotional affect of product involvement. *NA—Adv. Consum. Res.* **14**(March) (2010)
21. Mumcu, Y., Kimzan, H.S.: The effect of visual product aesthetics on consumers' price sensitivity. *Procedia Econ. Finan.* **26**(15), 528–534 (2015). [https://doi.org/10.1016/s2212-5671\(15\)00883-7](https://doi.org/10.1016/s2212-5671(15)00883-7)

22. Reimann, M., Zaichkowsky, J., Neuhaus, C., Bender, T., Weber, B.: Aesthetic package design: a behavioral, neural, and psychological investigation. *J. Consum. Psychol.* **20**(4), 431–441 (2010). <https://doi.org/10.1016/j.jcps.2010.06.009>
23. Seabrook, A.: What You Can Learn from Coca-Cola’s Marketing Strategy | Merca2.0. Retrieved November 25, 2019 (2019), from <https://www.merca20.com/what-you-can-learn-from-coca-colas-marketing-strategy/>
24. Tarta, C.: Studies and Articles Volume III Section: Economy And Management (November 2014) (2016)
25. Vaughn, C., Leff, J.: The interaction of life events and relatives’ expressed emotion in schizophrenia and depressive neurosis. *Br. J. Psychiatry* **136**(2), 146–153 (1980). <https://doi.org/10.1192/bjp.136.2.146>
26. Wang, F., Hannafin, M.J.: Technology-enhanced learning. *Environments* **53**(4), 5–23 (2005)
27. Wexner, L.B.: The degree to which colors (hues) are associated with mood-tones. *J. Appl. Psychol.* **38**(6), 432–435 (1954). <https://doi.org/10.1037/h0062181>
28. Zheng, X., Cheung, C.M.K., Lee, M.K.O., Liang, L.: Building brand loyalty through user engagement in online brand communities in social networking sites. *Inf. Technol. People* **28**(1), 90–106 (2015). <https://doi.org/10.1108/ITP-08-2013-0144>

The Brand-Cause Fit in the Advertising Campaign for Sprite's #YouAreNotAlone



Claudia Cachay-Marín , Francisco Arbaiza ,
and Eliana Gallardo-Echenique 

Abstract Several authors have determined concepts related to the inclusion of social problems in advertising. One of them is brand-cause fit, which refers to the condition in which a brand and a social problem—with which the target audience feels identified—are conceptually united in a single communication proposal. The purpose of this study was to analyze how university students perceive the brand-cause fit as a communication strategy in Sprite's "You Are Not Alone" campaign. A qualitative methodology was adopted and semi-structured interviews were conducted with 24 university students of which 12 belonged to the LGBT community and 12 were heterosexual. The participants positively perceived the brand-cause fit used as an advertising strategy in Sprite's campaign. Not only do they value the fact that brands include social issues and problems in their communication, but they also claim that it is the brands' responsibility to do so.

Keywords Advertising · Brand-cause fit · Social discourse · Sexual minority

1 Introduction

For some years now, it has been argued that traditional advertising -based on functional advantages and benefits- no longer works [1]. Consumers are no longer just looking for a product or service that meets their needs; what they crave most is to feel that the brand of such product or service understands and represents them [2]. Thus, new techniques such as cause-related marketing have emerged. This proposal allows consumers to contribute to society through a purchase, under the promise

C. Cachay-Marín · F. Arbaiza · E. Gallardo-Echenique (✉)
School of Communications, Universidad Peruana de Ciencias Aplicadas, Prolongación Primavera
2390, Lima 15023, Peru
e-mail: eliana.gallardo@upc.pe

C. Cachay-Marín
e-mail: u201312174@upc.edu.pe

F. Arbaiza
e-mail: francisco.arbaiza@upc.edu.pe

that the brand will give a certain amount of profits to a social cause [3]. Consumers have different interests in certain social causes such as the conservation of the planet, discrimination, or inequality; therefore, these type of campaigns continue to grow, managing to create a favorable situation for all stakeholders of the brand that promotes the cause [4]. This form of communication has become a constant for many brands. It is a strategy that not only increases sales, but also contributes to society [1, 5].

Contrary to the consumption proposal, social advertising arises as to a type of communication that conveys values and principles so that consumers accept, modify or abandon certain ideas or behaviors, and thus contribute to the improvement of social problems [6, 7]. Brand-cause fit is the concept in which a brand and a social “problem” are conceptually united to attain a new form of communication that creates a link with the target. The brand strategically selects the social problem, which is of interest to the target audience [8]. One reason why this form of communication is becoming an option is because there is a growing trend in advertising called brand responsibility, where the brand is linked to a social problem. Previous studies [8–10] have shown that the existing fit between the brand and the social problem has a great impact on the consumer.

Carrying out campaigns involving social issues is a form of advertising in which realities are represented and new ways of seeing problems are constructed while contributing to some social change [11]. Currently, several brands are choosing to implement social campaigns to create a favorable association in the consumer’s mind. However, choosing social discourse in campaign evidence certain social issues; therefore, the brand must be aware and responsible, as it can have an impact on the audience’s commitment to the issue and hence to the brand [11, 12]. Brand cause-fit refers to a brand and a social issue that come together to create a concept and advocate for this issue [8]. Thus, the brand takes a risk by deciding to engage in advertising with a social discourse [13]. It is likely that the brand will find it difficult to manage the balance between the marketing strategy and the social message; therefore, there should be a fit between the social issue chosen and the company’s mission and values [8].

Although brands are incorporating the support of social causes to improve their brand image into their advertising strategies, there is still little research on how these actions are affecting or influencing consumer decisions and attitudes towards them [14, 15]. In addition, it is important that the cause at hand is consistent with the brand-cause fit [4]. Creating campaigns with a social background is a trend created by consumers; but how much this type of advertising influences consumer attitudes is what needs to be researched [16]. Brands have power and influence in society and can contribute to change [16, 17]. Cause-related marketing is seen as an effective way to build the brand, increase sales, revitalize corporate values, create awareness, and make corporate social responsibility visible.

This study analyzes the perceptions of students at a private university towards a spot of Sprite’s ‘You Are Not Alone’ campaign launched in November 2019 [18]. This initiative proposed situations experienced by young people related to loneliness and insecurity, through the Reddit platform, where they told their feelings anonymously [18]. The brand then decided to extend the message to support the gay pride festival,

and precisely this is the campaign analyzed. Therefore, the question that guides this study is: How do university students perceive the brand-cause fit as a communication strategy in Sprite's 'You Are Not Alone' campaign?

2 Methodology

This study is situated in the interpretivist paradigm with a qualitative approach, to understand how people interpret various situations according to their experiences [19, 20]. Its design is phenomenological, where people describe some phenomenon according to their own experience [19], the brand-cause fit as an advertising resource, which touches on an important issue: the visibilization of the LGTB community. The data collection technique was the semi-structured interview, which is perceived as a purposeful conversation because, although it follows a question guide, it allows further questioning to go deeper and understand what the interviewee is referring to [21, 22]. An interview guide with 18 questions was prepared. Due to COVID-19, the interviews were conducted via video calls using the Google Meet application. The interview guide was validated in the field with the first interviewees [23]. Specifically, question number two (Have you ever seen a spot of any brand where social content is included?) was rephrased because some interviewees did not understand it.

A total of 24 university students from a private university participated in the study, 12 belonging to the LGBT community and 12 heterosexuals. Students from the different programs of the School of Communications were excluded since they are familiar with advertising terms, biasing the results. All participants were chosen through purposive sampling, in a non-random manner, as the sample is highly variable and not necessarily easily accessible [24]. This study took into account respect for others, so participants were not allowed discrimination of any kind. Verbal consent was obtained to record the interview, making it clear that everything said would be used only for the purposes of this study [25]. The recorded interviews were transcribed to facilitate the analysis process. Thematic analysis was chosen as the method for analyzing the interviews in order to identify key concepts into categories, or themes [26]. Each participant received an information sheet with the objectives and guidelines of the study [25] indicating that they could withdraw from the study without any prejudice; and, in order to respect their identity, each participant was identified by codes (i.e., P14, F, LGTB).

3 Results and Discussion

3.1 Brands Addressing Social Issues in Communication

Most of the interviewees recalled at least one advertisement that addressed some social issue or theme. However, many of the examples given did not belong to Peruvian brands. They also mentioned that the Peruvian public is still very sensitive to certain social issues and that brands do not dare to include them in their communication. Brands have become more inclusive in terms of their content, but this inclusion can lead to negative effects, especially in audiences with opposing positions; therefore, brands often do not dare to take the final step for fear of their audience's negative reaction [27].

The reaction to brands that touch on social issues was positive; however, almost all agreed that some brands address them influenced by “fashion,” in order to have an inclusive brand image that is concerned about society, when in fact it is not true. The response regarding the use of social causes came mostly from participants belonging to the LGBT community. The students mentioned that advertising does have a responsibility to society due to its power, that it is not enough to show the needs of the individual to sell a product/service, but that it has to show the real context in which he/she lives to be more credible and real. Furthermore, the power of advertising lies not only in its persuasive capacity but also in its presence in people's lives and in society [28]. Therefore, the existing relationship between social discourse and advertising lies in advertising discourse harnessing the power it has for the change and improvement of societies [29, 30].

I sincerely believe, being from the LGBT community, that the community is being used a lot as a business, as something to sell more, as something to say “Hey, we support gays and lesbians” (...) And to think that the LGBT community does not really need a commercial, it needs to be defended, it needs to be more listened to and erase so many stereotypes about them. (P14, F, LGTB)

I think advertising is a very interesting tool (...). We see it today in our phones, in the streets. Nowadays, (...) advertising is a very useful tool not only to offer a product, to offer a service, but it is a platform for many messages or some kind of interest that a company wants or needs to show. (P11, M, LGTB)

Participants recognize that there are differences between the social issues addressed in mass media (television, radio) and digital media (social networks) in that sensitive issues are addressed more frequently in virtual platforms than in broadcast media. One possible explanation is that Peruvians are still very sensitive to certain issues. Although brands are afraid to address inclusion because of possible reactions from their audiences, they have reconsidered their advertising content and have begun to show an image with inclusive messages for mass audiences and no longer for a niche market [31–33].

3.2 *The Social Issue Chosen by the Brand*

Interviewees stated that it is important for brands or companies to have a clear position in the social sphere, and to raise their voice when necessary, which is still not entirely the case in Peru. It is important that there is a logical relationship between what the brand says and what the brand is, otherwise consumers might think that they only use certain social issues to profit from them and obtain more sales. The fit between the social cause and the brand, where both should be consistent, can positively impact brand equity, while a low-fit can be inconsistent with and distracting from the brand equity as well as affect the consumer's perceived value of the brand [4]. Additionally, the fit between the social issue and the brand may impact both the final purchase decision and the attitude towards the campaign [10, 34].

There is a lot of posing, which is a bit—how can I say it—more “performative.” What's more, they give an image that they do not necessarily share, saying that they are socially responsible when they are not. (P08, M, No LGTB)

There does need to be a relationship between what they show in their communication campaigns and the actions they themselves take at the organizational level or as part of social responsibility, because otherwise the message they give is something banal or that they only do to go along with “social fashions” (...) Sometimes I don't know if they really want to make the issue visible or if they just do it to sell. (P15, M, LGTB)

To determine whether the social issue is addressed to achieve a change, the message should not be transitory and appear at convenient times, but on the contrary, it should be constant. The efficiency of the effect of advertising linked with social issues will depend on whether the brand focuses its attention on social issues that need to be exposed [35]. Advertising has a social responsibility that implies the need for ethics on the part of advertisers and publicists [2, 29].

For example, with movements such as for gay pride month or LGTB or racial campaigns, they show a lot of involvement, but as soon as the moment passes, “bam”! they forget. For just one month they care about LGBT people or people of color and then there is nothing. (P05, F, LGTB)

On the one hand, participants showed a positive attitude towards brands that address social issues only if the communication is genuine, constant, and risky; but, on the other hand, the attitude was confused and offensive if there is no relationship between the brand and social issues. There has to be a logical relationship (fit) in line with the brand image between the social cause and the essence or perception that consumers have towards the brand to obtain a positive attitude from users [3, 15]. The brand-cause fit uses a social issue in such a way that its audience is reflected and identifies itself, thus achieving effective communication [8].

I would not value it if I see that the message they are trying to bring across with the communication campaign is not accompanied by the company's behavior (...) otherwise for me it is a job done halfway. (P16, M, No LGTB)

It is disrespectful for people who are looking for a change regarding this social problem that a company only profits and does not contribute to anything —quote unquote— positive (...) Because for an advertisement to help a social problem, this advertisement needs to show the background of the problem and not simply sell the product. (P21, M, LGTB)

3.3 Brands No Longer just Tell, They also Do

After watching Sprite's 'You Are Not Alone' campaign, participants expressed various positive feelings and identification with the LGBT community. For heterosexual students, in addition to generating emotion, it also generated empathy to recognize the privileges they have by not being part of a group repressed by society. Although the spot is taken positively, participants agree that the perception they have of the brand has neither improved nor worsened, but has remained the same. The reason is that these types of issues should be common and "normal" in any communication. The power of advertising should not only be reflected in the commercial aspect; hence, there are different social initiatives in advertising, such as the exhibition of relevant social issues and even the intention of educating people, which can open important and necessary debates. Advertising messages have not only helped social causes, but also initiatives to achieve change [11, 30].

They generate satisfaction, happiness because, finally, when you see content that goes according to your convictions, to what you believe in, you feel happy because you have found an organization, a brand that reflects your identity, what you think and, in a certain way, this satisfies you. (P04, F, LGTB)

Sprite's campaigns are really a bit more relaxed compared to Coca-Cola's, for example, which sells you more the "Open Happiness" and all those things; more active, more youthful. They are aimed at young people, love, fun —what do I know— and I think it is in line with what Sprite does. (P13, M, No LGTB)

For the interviewees, Sprite is a globally recognized and positioned brand, so it can address sensitive and relevant issues; however, they believe that this would change if the brand were small or not so recognized. Moreover, the fact that the product is not the main focus of the spot was highly valued; this generated confidence and a feeling that the communication is genuinely made to achieve a change or social awareness, as it is not perceived that the ultimate goal is to obtain financial gains. Advertising cannot necessarily radically change behaviors by making a spot, but it can put social issues on the table and that can make a big difference given that consumers now prefer brands [31] that use inclusive messages for mass audiences, leaving the idea that they can only be made for a niche market behind [32].

Throughout the entire spot, at no time did I see anything that lasted more than five seconds that was purely Sprite. It focused on what the person is like and that's when you say: How cool that the brand has taken the time and resources to do something as important as this! (P17, F, LGTB)

I don't know if such a short spot can change anyone's mentality —I would like to think so—, but it seems to me that at least seeing this spot might make other people curious, and they might ask and find out more about it. (P02, F, LGTB)

4 Conclusions

How do university students perceive the brand-cause fit as a communication strategy in Sprite's 'You Are Not Alone' campaign? The university students positively

perceived the brand-cause fit used as an advertising strategy in Sprite's campaign. Not only do they value the fact that brands include social issues and problems in their communication, but they also claim that it is the brands' responsibility to do so. However, they are aware that there are many brands that take advantage of certain social issues just to reach their audience, without having a real intention to help society. The students mentioned two key factors to be able to differentiate Sprite as a genuine brand from one that is not: (a) the message must be permanent and not only for specific times; and (b) the brand must not only show but defend the cause or social issue.

Regarding the fit that should exist between the brand and the social issue, there needs to be consistency between what the brand says and what the brand is, otherwise it would not only lose credibility, but would also generate a negative attitude. The attitude towards Sprite's campaign was positive because the brand is known for its involvement in social issues. However, this would not influence the participants' purchase decision, as they consider that addressing social issues should be something normal. The literature states that in brand cause-fit the issue should be chosen for the identification it has with its audience, but this study showed the opposite. Sprite's spot was aimed at the LGTB community, but both participants from the LGTB community and heterosexuals had the same positive reaction to the spot and to the brand. This shows that what is important is not the user's identification with the social theme or issue, but that the brand must be genuine, consistent, responsible, and above all honest when communicating these issues.

This study has two limitations. On the one hand, the study sampled a specific group of students from a private university, so the findings cannot be generalized, but they are transferable to participants with the same characteristics. On the other hand, it was conducted during the COVID19 pandemic and, because of it, the interviews were conducted online and most of the participants decided not to turn on their camera; therefore, it was not possible to have the same interaction as with a face-to-face interview. For future lines of research, the sample should be extended by widening the age range, having as participants not only students but also people with different occupations to analyze other advertising campaigns of other companies and/or brands that involve other social issues that are sensitive for Peruvians.

References

1. Champlin, Y., Sterbank, S.: Agencies as agents of change: considering social responsibility in the advertising curriculum. *J. Advert. Educ.* **22**, 137–143 (2018). <https://doi.org/10.1177/1098048218807138>
2. Barragan, M., Flores, E.: La publicidad en la Mercadotecnia social advertising in the social marketing. *Int. J. Good Conscienc.* **14**, 231–234 (2019)
3. Rohit, D., Panda, R.: Cause-brand association and consumer attitude: a review. *J. Manag. Res.* **18**, 13–26 (2018). <https://doi.org/10.18843/ijcms/v9i1/09>
4. Handa, M., Gupta, S.: Digital cause-related marketing campaigns. *J. Indian Bus. Res.* **12**, 63–78 (2020). <https://doi.org/10.1108/JIBR-09-2019-0285>

5. Natarajan, T., Jublee, D.I., Kasilingam, D.L., Stephen, G.: The moderating role of social themes in cause-related marketing advertisements. *Int. Rev. Public Nonprofit Mark.* **15**, 433–454 (2018). <https://doi.org/10.1007/s12208-018-0208-4>
6. Martínez, M., Domínguez, M., Sotelo, G.: “No me ignores”: Publicidad social que aborda el tema de los feminicidios. **11**, 1683–1688 (2019)
7. Senes, F., Ricciulli-Duarte, D.: La publicidad social como herramienta para la construcción de ciudadanía. *Razón Crítica.* 183–199 (2019). <https://doi.org/10.21789/25007807.1425>
8. Champlin, S., Sterbenk, Y., Windels, K., Poteet, M.: How brand-cause fit shapes real world advertising messages: a qualitative exploration of ‘femvertising.’ *Int. J. Advertising.* **38**, 1240–1263 (2019). <https://doi.org/10.1080/02650487.2019.1615294>
9. Pracejus, J., Olsen, D.: The role of brand/cause fit in the effectiveness of cause-related marketing campaigns. *J. Bus. Res.* **57**, 635 (2004). [https://doi.org/10.1016/S0148-2963\(02\)00306-5](https://doi.org/10.1016/S0148-2963(02)00306-5)
10. Pracejus, J.W., Deng, Q. (Claire), Olsen, G.D., Messinger, P.R.: Fit in cause-related marketing: an integrative retrospective. *J. Glob. Sch. Mark. Sci.* **30**, 105–114 (2020). <https://doi.org/10.1080/21639159.2018.1509365>
11. Alcaide-Lara, E.R.: Imagen social y contextos socioculturales en el discurso publicitario institucional español con fines sociales. *Pragmática Sociocult./Sociocult. Pragmat.* **7**, 297–334 (2020). <https://doi.org/10.1515/soprag-2019-0026>
12. Kim, K., Cheong, Y., Lim, J.: Choosing the right message for the right cause in social cause advertising: type of social cause message, perceived company—cause fit and the persuasiveness of communication. *Int. J. Advertising.* **34**, 473–494 (2015). <https://doi.org/10.1080/02650487.2015.1006081>
13. Bigné-Alcañiz, E., Currás-Pérez, R., Ruiz-Mafé, C., Sanz-Blas, S.: Cause-related marketing influence on consumer responses: the moderating effect of cause-brand fit. *J. Mark. Commun.* **18**, 265–283 (2012). <https://doi.org/10.1080/13527266.2010.521358>
14. Huertas-García, J., Lengler, C., Consolación-Segura, R.: Co-branding strategy in cause-related advertising: the fit between brand and cause. *J. Prod. Brand Manag.* **26**, 135–150 (2017). <https://doi.org/10.1108/JPBM-07-2015-0939>
15. Buil, I., Martínez, E., Montaner, T.: La influencia de las acciones de marketing con causa en la actitud hacia la marca. *Cuad. Econ. y Dir. la Empres.* **15**, 84–93 (2012). <https://doi.org/10.1016/j.cede.2012.01.002>
16. Shoenberger, H., Kim, E., Johnson, E.: Role of perceived authenticity of digital enhancement of model advertising images on brand attitudes. *Soc. Media Engagement J. Interact. Advert.* **20**, 181–195 (2020). <https://doi.org/10.1080/15252019.2020.1840459>
17. Mamuric, N.: Femvertising: advertising taking social responsibility to sell brands. *J. Brand Strateg.* **7**, 318–325 (2019)
18. Mercado Negro: “No estás solx”, la continuación de la campaña de Sprite que se roba todos los elogios, <https://www.mercadonegro.pe/publicidad/campanas/no-estas-solx-la-continuation-de-la-campana-de-sprite-que-se-roba-todos-los-elogios/>, (2019).
19. Merriam, S., Tisdalle, E.: *Qualitative Research. A Guide to Design and Implementation.* Wiley (2014)
20. Creswell, J.W.: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* SAGE Publications Inc., Thousand Oaks, CA (2014)
21. Glesne, C.: *Becoming Qualitative Researchers: An Introduction.* Pearson (2006). <https://doi.org/10.3316/qrj0602203>
22. Patten, M., Newhart, M.: *Understanding Research Methods: An Overview of the Essentials.* Routledge (2017)
23. Namakforoosh, M.N.: *Metodología de la investigación.* Editorial Limusa S.A. Grupo Noriega, Mexico, D.F. (2013)
24. Mendieta-Izquierdo, G.: Informantes y muestreo en investigación cualitativa. *Investig. Andin.* **17**, 1148–1150 (2015). <https://doi.org/10.33132/01248146.65>.
25. Rapley, T.: *Los análisis de la conversación, del discurso y de documentos en Investigación Cualitativa.* Ediciones Morata, S.L., Madrid, España (2014)

26. Nowell, L.S., Norris, J.M., White, D.E., Moules, N.J.: Thematic analysis: striving to meet the trustworthiness criteria. *Int. J. Qual. Methods* **16**, 1–13 (2017). <https://doi.org/10.1177/1609406917733847>
27. Martínez, C., Montiel, Ó., Rodríguez, C., Tovar, J.: Real Inclusion or Simulation? An exploration of an Advertising Campaign for the Gay Community in Mexico. *Rev. Telos*. **22**, 204–222 (2020). <https://doi.org/10.36390/telos221.14>
28. Kendrick, A., Fullerton, J.A., Kim, Y.J.: Social responsibility in advertising: a marketing communications student perspective. *J. Mark. Educ.* **35**, 141–154 (2013). <https://doi.org/10.1177/0273475313490171>
29. Pellicer, M.: Advertising and its social responsibility. *Rev. Comun. Vivat Acad.* **19**, 43–51 (2017). <https://doi.org/10.15178/va.2017.139.43-51>
30. Alvarado Lopez, M.C.: La publicidad en el marco de la comunicación para el desarrollo: hacia un nuevo modelo de publicidad para el cambio social. *CIC Cuad. Inf. y Comun.* **17**, 191–207 (2012). https://doi.org/10.5209/rev_CIYC.2012.v17.39264
31. Cunningham, G., Melton, N.: Signals and Cues: LGTB inclusive advertising and consumer attraction. *Sport Mark. Q.* **23** (2014)
32. Eisend, M., Hermann, E.: Consumer responses to homosexual imagery in advertising: a meta-analysis. *J. Advert.* **48**, 380–400 (2019). <https://doi.org/10.1080/00913367.2019.1628676>
33. Northey, G., Dolan, R., Etheridge, J., Septianto, F., Van Esch, P.: LGBTQ imagery in advertising how viewers political ideology shapes their emotional response to gender and sexuality in advertisements. *J. Advert. Res.* **60**, 222–236 (2020). <https://doi.org/10.2501/JAR-2020-009>
34. Cheron, E., Kohlbacher, F., Kusuma, K.: The effects of brand-cause fit and campaign duration on consumer perception of cause-related marketing in Japan. *J. Consum. Mark.* **29**, 357–368 (2012). <https://doi.org/10.1108/07363761211247479>
35. Suprun, V.M., Suprun, L.V.: Social advertising in modern communication space. *Sci. Innov.* **14**, 62–68 (2018). <https://doi.org/10.15407/scine14.04.062>

Consumer Bullying in Online Brand Communities—Quantifying a Dark Social Media Phenomenon



Oliver Warke, Jan Breitsohl, and Joemon Jose

Abstract Whilst authors in the marketing literature have made some qualitative suggestions that aggressive consumer-to-consumer interactions (i.e., bullying) negatively affect users' social media experience, there is no quantitative evidence to substantiate such propositions. Moreover, several studies emphasize that large brands such as Nike and Coca Cola are reluctant to get involved when consumer interactions turn hostile. Applying text mining techniques from computer science to a novel phenomenon in marketing, this paper aims to provide a quantitative overview of the prevalence and impact of bullying in online brand communities, and the potential effectiveness of brand responses. To this purpose, we scraped data from the official Facebook community of 14 retail brands (e.g., Lidl, Tesco) for a 3-month period, analyzed the content and sentiment of consumer comments that classified as bullying, and ran t-tests to measure the impact of a brand response. Analyzing a total of 49,866 comments, our findings offer detailed insights on the average bullying rate for retailers' online brand communities and significant variations in bullying between these retailers. Furthermore, we show that bullying has a largely negative emotional impact on community users, and that a brand who responds to bullying can cause a substantial, positive sentiment shift. We thus offer theoretical and managerial contributions to the novel and under-researched 'dark side of social media' in the digital marketing literature.

Keywords Bullying · Online communities · Brands · Moderation · Management

O. Warke · J. Jose

School of Computing Science, University of Glasgow, 18 Lilybank Gardens, Glasgow G12 8QQ, Scotland, UK

e-mail: Oliver.Warke@glasgow.ac.uk

J. Jose

e-mail: Joemon.Jose@glasgow.ac.uk

J. Breitsohl (✉)

Adam Smith Business School, University of Glasgow, Gilbert Scott Building, Glasgow G12 8QQ, Scotland, UK

e-mail: Jan.Breitsohl@glasgow.ac.uk

1 Introduction

Social media platforms such as Facebook have become an increasingly dominant means for consumers to communicate with each other. However, market research shows that comments on social media have become increasingly hostile, with over 50% of users experiencing different forms of aggression on a regular basis in both the UK and the US [13, 24]. Aggressive communication behavior carried out on social media is generally referred to as cyberbullying [20], and we use this term to encapsulate similar online phenomena such as hate speech, incivility or trolling. Meta-analytic studies evidently show that being the victim of as well as merely witnessing cyberbullying reduces one’s life satisfaction, emotional security, and performance in daily tasks [25], and scholars have described it as one of the key social challenges of the twenty-first century [19]. Increasingly, cyberbullying takes places in social media communities that are hosted by brands, hereafter termed ‘online brand communities’. Take, for instance, the example from the H&M brand community on Facebook in Fig. 1, on which consumers engage in an increasingly aggressive debate about whether the brand promotes racism.

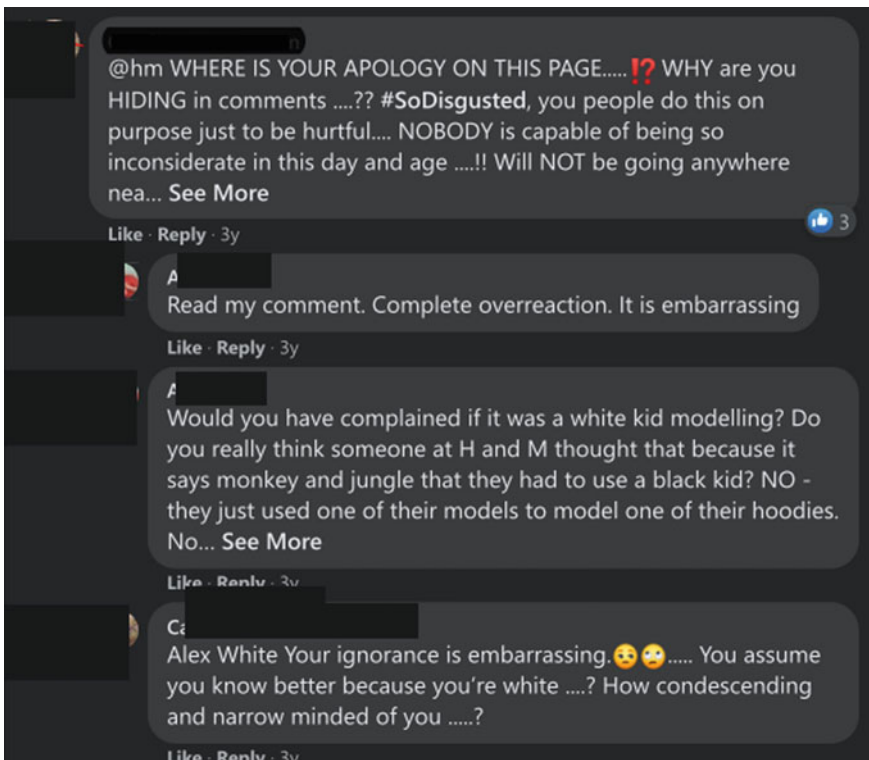


Fig. 1 Example of cyberbullying on H&M online brand community on Facebook

Surprisingly, the marketing literature has so far largely ignored cyberbullying behaviors and instead focused on the positive aspects of consumers' engagement in online brand communities [11]. Authors highlight, for instance that such communities offer consumers a number of social and functional benefits, whilst corporate brand owners are able to learn about and interact with current and future target groups [5]. Evidence of 'dark behaviors' such as cyberbullying is scarce, and relies on online observations and qualitative analyses, offering some insights on its psychological and social underpinnings [15]; however, quantitative insights are largely missing and have been called for [10]. For instance, Breitsohl et al. [6] use online observations and manual thematic analysis to identify that consumers engage in different types of cyberbullying, and call for complementary research that can offer insights on the prevalence of cyberbullying across industries. Similarly, Dineva et al. [12] uses online observations to uncover that brands largely do not respond to incidents of cyberbullying in their official communities on Facebook, and they call for research to provide evidence on both the negative social impact of such incidents as well as a quantification of brand response behaviors. Likewise, several recent reviews on the state of the digital marketing literature generally emphasize that more research is needed in relation to negative interactions between consumers on social media [21, 29, 35].

These calls for research motivate the present paper. Using a combination of text mining techniques, we captured 49,866 comments over a period of three months from the official Facebook brand communities of 14 UK and US retail brands (e.g., McDonald's, Tesco). We used a combination of sentiment analyses, content analyses, and statistical tests to identify the prevalence of cyberbullying, to measure its emotional impact and to test whether a brand response makes a difference to how consumers feel. Overall, our findings show that brands experience an overall bullying rate (i.e., percentage of total comments that were classified as bullying) of 10.1%, ranging from 50.85% for the retailer with the highest rate to 0.27% for the retailer with the lowest rate. Moreover, we evidence that the impact of bullying includes an increase in negative emotions of 5.61% and a decrease in positive emotions by 4.55%. Finally, a *t*-test reveals that a response from a brand to a cyberbullying incident leads to more positive sentiments compared to when the brand does not respond.

To the marketing literature, our findings contribute ecological validity to the conceptualization of cyberbullying in online brand communities. Marketing researchers increasingly highlight the importance of underlining theoretical propositions with real-world evidence [33], and we complement previous qualitative propositions with big data insights on the prevalence of cyberbullying. We also offer insights on the emotional impact of cyberbullying in online brand communities, an area which has rarely been looked at and is largely limited to a positive–negative valence [18], whilst we unpick a set of distinct emotions such as anger and sadness.

To the technology and computer science literature, we further offer a methodological contribution by combining three classifiers to improve the reliability and precision of our data. Prior research usually relied on one classifier to capture different forms of cyberbullying [8], whilst we highlight the utility of combining three different classifiers that capture different nuances of cyberbullying.

2 Literature Background

Cyberbullying describes any behavior performed through electronic or digital media by individuals or groups that repeatedly communicates hostile or aggressive messages intended to inflict harm or discomfort on others [31]. Early studies on cyberbullying were pre-dominantly conducted in relation to private online interactions via peer-to-peer messaging or emails [16]. However, with the constant rise of public groups on social media, the scholarly focus has shifted to cyberbullying that occurs on communal platforms, including forums, blogs and social media [34]. Scholars further note that to qualify as cyberbullying, the elements of repetition and intention do not need to be present when cyberbullying occurs on social media platforms. Slonje et al. [28] suggest that messages on social media remain permanent, and as such can be read again and again by both the victim and an audience. Moreover, the heterogeneity of publicly accessible online communities renders individual's ability to understand the intent behind a message an increasingly difficult undertaking [5]; unlike interactions on more private channels such as emails, where those involved tend to know each other and usually have a shared background, cyberbullying in online communities tends to occur between strangers from different backgrounds, making it difficult to decipher whether a message was meant to be harmful, and therefore constitutes cyberbullying [4].

According to the psychological literature, cyberbullying can occur in a number of different contexts, including interactions about gender issues, race, ethnicity, sexual orientation, political opinions, and physical attributes [16]. The consequences of being cyber-bullied range from lower life satisfaction, emotional security, and performance in daily tasks [25] to lower emotional security and lower trust in social interactions [31]. Meta-analytical evidence suggests that these consequences depend upon the form that cyberbullying takes as well as how witnesses to a cyberbullying interaction react [19]. Different forms of cyberbullying lead to different outcomes; a comment that uses humor (e.g., teasing) tends to be seen as less severe than comments that use swearwords and no humor (e.g., harassment) [34]. Other than looking at consequences, some authors have started to explore how authorities with some form of social power may respond to cyberbullying. Free-hand approaches are rejected due to the consensus that cyberbullying is harmful and needs to be prevented [19]. Likewise, researchers rarely include censorship in studies on cyberbullying interventions, perhaps because it does not prevent the bully to repeat his/her behavior and fails to offer victims explicit (public) support [22]. However, based on research in related fields, there seems to be a general agreement that Internet users support freedom of speech and disapprove of censorship.

In comparison, consumer cyberbullying differs in three important ways. First, as outlined by Breitsohl et al. [6], consumer cyberbullying takes place between followers of brands that is in a consumption-driven context. Marketing research highlights the importance of brands as an identity-central factor in contemporary (consumer) societies [32], and the identification with brands in online communities can lead to aggressive interactions [15]. Yet, whilst cyberbullying in the context of brands may

lead to comparable interactions than those based on issues of race or ideology, the psychological literature has so far not considered cyberbullying in a brand context. A first indication that C2C interactions may differ comes from Breitsohl et al.'s [6] study, identifying six types of cyberbullying, three of which have not been reported in cyberbullying research before. Specifically, the authors highlight that similar to other cyberbullying contexts, consumer in online brand communities engage in teasing, criticism, trolling and harassment, and they further identify two novel types of cyberbullying, ostracism (ignoring other consumers) and camouflage (ambiguous messages).

Second, the host of an online brand community (i.e., the corporation that owns the brand) has a commercial interest in those involved in consumer cyberbullying (i.e., its brand followers), which is likely to affect the governance mechanisms [26]. Existing work in cyberbullying concentrates on the perspective of parents, teachers, legislators and independent, non-commercial forum hosts [34], all of which are in non-commercial relationships with the parties involved in cyberbullying. In contrast, brand followers represent actual or potential customers that represent commercial value to the corporate community host. Consequently, in fear of losing customers, corporate brand owners will be less likely to engage in authoritative interventions that punish, or lecture bullies as is common to other forms of cyberbullying. Dineva et al. [12] for instance shows that 5 out of 6 corporate brand owners hosting an online brand community choose not to respond to aggressive consumer interactions at all, whilst such non-intervention is not considered in the psychological literature. Similarly, brand followers are likely to be aware of the commercial power they hold over corporate brand owners [11], and this may be reflected in their reactions toward corporate interventions.

In sum, consumer cyberbullying is unique in its consumer brand focused interaction context, its commercial consequences, and its corporate host (i.e., the brand). However, marketing knowledge to this regard is limited, and there is little large-scale quantitative evidence on its prevalence, how it affects other consumers, and whether a brand response can be impactful.

3 Methodology

We implemented a data crawler to collect consumer comments. Fourteen international retail brands were selected, including Tesco, Asda, Aldi, Lidl, Morrisons, and Marks and Spencer's, Primark, Dr Martens, and Next, Costa Coffee, KFC, Subway, Greggs, and McDonald's. Python was chosen as the language to use due to its fast deployment, ease of use, and a vast array of supporting libraries. We decided to use the Facebook mobile site instead of the regular Facebook site since data can be accessed more readily. To navigate the site and load comments dynamically, Selenium, a web automation tool, was employed. Selenium allows the browser to mimic a human's actions, such as entering search queries and pressing buttons. We then used BeautifulSoup to analyze the downloaded page. BeautifulSoup is a library

designed to help get data from a webpage's HTML code and helps cleaning the data from the webpage before identifying its elements. As each comment was processed, data was then inserted into the MongoDB database.

We then employed a set of classifiers. Classification involves implementing machine learning algorithms [7] to train classification models. These models group input data into a set of predefined groups, predicting to which group each given piece of input data belongs. For this paper, we used sentiment and emotional classifiers. Sentiment classifiers frequently group input data into three categories; positive, negative, or neutral (neither) by producing a sentiment score using the classification model [23]. Emotional classifiers can categorize data as a variety of emotions such as happiness, sadness, anger, and fear [27]. In lack of applicable cyberbullying classifiers, we used the hate speech and offensive language classifier from Davidson et al. [9]. The classifier can categorize text as 'Hate Speech', 'Offensive Language', or 'Neither', has proven a high reliability in recent research and was used by others in the past to identify cyberbullying [14].

We further used NLTK sentiment analysis with VADER [17] combined with two classifiers. First, the NRC Lexicon python package by Bailey [3], a crowd-sourced lexicon with manually assigned annotations for each word which makes use of the NRC Word-Emotion Association Lexicon and uses approximately 27,000 words. Second, to capture the constantly growing use of emojis [1], the text2emotion classifier [30] uses a custom lexicon and the NLTK library alongside the emoji library to assign emotions to emojis, allowing their inclusion in the classification of a piece of text. The NLTK library further allows for tokenization and word lemmatizing. We used the following definition for a comment to be deemed as cyberbullying:

A comment chain containing a comment classified as having a significant negative emotion by one of the emotional classifiers or classified as hate speech by the third classifier or classified as offensive language with strong negative sentiment. The comment chain would then have to contain another comment by a different consumer which met the same criteria. This definition is in accordance with suggestions by Breitsohl et al. [6], describing consumer cyberbullying as a social media interaction process in which one consumer makes a potentially offending comment toward another consumer; 'potentially offending' includes comments that contain any form of interpersonal aggression (e.g., hate speech, insults), but also non-aggressive disagreements (e.g., constructive criticism, light teasing).

4 Results

Overall, our descriptive statistics show a total of 5035 bullying instances from 49,866 comments (10.1%). Bullying rates had a range from 0.27 to 51.85%. However, the percentage varies widely between brands. Interestingly, KFC recorded the largest bullying rate (51.85%), followed by Sainsbury's (45.04%). The mean across the brands was 15.13%, and the median was 8.82%. These results are illustrated in Table 1.

Table 1 Comment totals and percentage rates

Brand	Brand comments (totals)	Bullying comments (totals)	All captured comments	Bullying rate (%)	Brand response rate (%)
McDonalds	45	629	8634	7.29	0.52
Morrisons	31	906	10,641	8.51	0.29
Tesco	116	449	3556	12.63	3.26
Aldi	113	622	8092	7.69	1.40
Lidl	84	331	3222	10.27	2.61
Costa Coffee	63	647	3043	21.26	2.07
KFC	0	70	135	51.85	0
Subway	3	206	1096	18.80	0.27
Sainsbury’s	76	499	1108	45.04	6.86
Primark	108	2	730	0.27	14.8
Greggs	6	52	587	8.86	1.02
Next	5	37	422	8.77	1.18
M & S	41	560	7876	7.11	0.52
Dr. Martens	28	25	724	3.45	3.87

The results presented in Table 2 show the range of emotions experienced by consumers in response to bullying, in the form of the percentage of the total emotions experienced that a particular emotion contributes. The full range of detected emotions consisted of those grouped as negative (anger, fear, disgust, and sadness) and positive (trust and joy). The results from the analysis show that all emotions within the negative emotions group occur at a higher rate (5.61% higher) within bullying instances than in non-bullying instances, and all emotions in the positive emotions group occur at a lower rate (4.55% lower) within bullying instances than in non-bullying instances. Interestingly, trust experiences the largest score for both bullying and non-bullying.

Table 3 shows the mean sentiment score observed across each community. The Vader sentiment intensity analyzer was used to obtain these scores and scores sentiment on a range of -1 to 1, with -1 being extremely negative and 1 being extremely positive. To test the effect of a brand response on consumer sentiment, we ran a two-sample *t*-test (two-tailed). Our findings indicate that it makes a significant difference ($t(602) = 5.6, p < 0.001$) in relation to consumers’ sentiment when comparing bullying interactions to which a brand responds ($M = 0.21, SD 0.5$) with bullying interactions to which a brand does not respond ($M = 0.08, SD 0.45$); in other words,

Table 2 Comparison of emotions after bullying versus no bullying comment chains

	Surprise	Fear	Sadness	Disgust	Anger	Trust	Joy
Bullying	5.04	6.71	6.61	4.51	5.92	13.36	9.75
Non bullying	5.95	5.73	5.9	3.92	4.74	14.95	11.64

Table 3 Mean Vader sentiment score observed on each brand page across comment chains with and without a brand response

Brand	No brand response	Brand response
Morrisons	0.04	0.08
Aldi	0.14	0.18
M&S	0.06	0.2
Dr. Martens	0.18	0.28
Next	0.03	0.11
Greggs	0.14	0
Sainsbury's	-0.04	0
Costa Coffee	0.02	-0.06
McDonalds	-0.04	0.16
Tesco	0.06	0.23
Lidl	0.11	0.09

when a brand responds, the consumer sentiment is more positive. It is important to note that whilst the sentiment analysis shows a larger positive difference in the mean compound score when the brand was present (0.21) than when it was not present (0.08), it does not indicate an overall shift of negative comments to positive comments; rather, it signifies a consistent increase in the mean comment sentiment after a brand intervention. Table 3 illustrates the sentiment shift per brand. Whilst the majority of supermarket brands had a positive sentiment shift, there are significant variations. For example, Tesco had a sentiment shift of 0.17, whilst Aldi had a shift of 0.04. Presumably, this has to do with the content of the brand response itself.

5 Discussion and Conclusion

This paper set out to offer first quantitative insights on the prevalence and impact of consumer brand bullying in online brand communities, as well as the usefulness of an intervening response from the brand. Our results substantiate the ecological relevance of this new phenomenon, indicating that 10.1% of comments in online brand communities relate to an undesirable consumer interaction behavior. Whilst prior studies had provided initial evidence on its existence [9, 15], we are first to offer quantitative evidence on the extent to which cyberbullying permeates communities, thus complementing and contributing to its theoretical relevance. Interestingly, bullying rates differed between different retailers, and future research should explore the reasons for these differences. For instance, do some brand systematically censor all negative consumer comments so that we did not observe anything—as suggested by Stephen [29] for instance—or are the consumers of a particular retailer less likely to engage in bullying because of some moderating characteristics (e.g., social class, reputation of retailer) as pointed out by Dineva et al. [12, 2]? Practitioners can benefit

from comparing their brand to others in the industry, and perhaps consider if their customer base needs explicit rules of interaction when bullying rates are high.

Our findings also contribute to the literature by offering empirical evidence on prior suggestions that cyberbullying has negative consequences. In the marketing literature, studies have suggested that hostile comments can have positive results (e.g., attracting more comments and increasing engagement; [18]), whilst others suggest the opposite [2], yet without being able to verify this perspective. Our results that cyberbullying increases negative emotions, and decreases positive emotions, indicates that on a large-scale, consumers do not seem to receive these comments positively. This raises further questions as to whether specific emotions translate into subsequent negative behaviors such as exiting the community [12]. Moreover, researchers are encouraged to investigate why some emotions are stronger than others (here: trust), and if this is consistent for both bullying and non-bullying. For instance, Breitsohl et al. [6] suggest that some forms of bullying can be more ambiguous and lead to emotional confusion. In any case, our result suggests that practitioners need to acknowledge, and address, cyberbullying as a consumer behavior that negatively impacts how their brand community users feel.

Finally, authors in the marketing literature who had observed that most brands do not actively intervene in cyberbullying have called for quantitative evidence on brand response effectiveness [2, 12]. Our finding that in most cases, consumer sentiment is significantly more positive when a brand responds to cyberbullying addresses such calls, and highlights the importance to investigate brand response strategies in future research. In particular, given that some brands' responses (e.g., Tesco) are able to cause much larger sentiment shifts than others (e.g., Aldi), investigating specific response characteristics (wording, tone, timing) that can explain these differences are a fruitful area of future research. Recent studies in a non-profit context already indicate that brands may make use of a variety of strategies [10] that if effective, improve an organization's image [11]. For social media content managers, our findings encourage joint, industry-wide efforts to learn from one another and co-develop effective response messages.

To conclude, our paper provides a first step toward bringing cyberbullying to the attention of marketing researchers and practitioners alike. Given that 10.1% of comments classify as cyberbullying, and that a brand who responds can have a positive emotional impact on its consumers, this study contributes to a new stream of research to tackle one of the key social challenges of the twenty-first centuries [19] within the realms of digital marketing practice.

References

1. Alshenqeeti, H.: Are emojis creating a new or old visual language for new generations? An asocio-semiotic study. *Adv. Lang. Literary Stud.* 7 (6) (2016)
2. Bacile, T.J., Wolter, J.S., Allen, A.M., Xu, P.: The effects of online incivility and consumer-to-consumer interactional justice on complainants, observers, and service providers during social

- media service recovery. *J. Interact. Mark.* **44**, 60–81 (2018)
3. Bailey, M.M.: NrcLex. <https://github.com/metalcorebear/NRCLEX> (2019). Last accessed 2021.09.01
 4. Bollmer, J.M., Harris, M.J., Milich, R., Georgesen, J.C.: Taking offense: effects of personality and teasing history on behavioral and emotional reactions to teasing. *J. Pers.* **71**(4), 557–603 (2003)
 5. Breitsohl, J., Kunz, W.H., Dowell, D.: Does the host match the content? A taxonomical update on online consumption communities. *J. Mark. Manag.* **31**(9–10), 1040–1064 (2015)
 6. Breitsohl, J., Roschk, H., Feyertag, C.: Consumer brand bullying behaviour in online communities of service firms. In: *Service Business Development*, pp. 289–312. Springer (2018)
 7. Brew, C.: Language processing: statistical methods. In: Brown, K. (ed.) *Encyclopedia of Language & Linguistics (Second Edition)*, 2nd edn, pp. 597–604. Elsevier, Oxford (2006). <https://www.sciencedirect.com/science/article/pii/B0080448542009482>
 8. Dani, H., Li, J., Liu, H.: Sentiment informed cyberbullying detection in social media. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 52–67. Springer, Cham
 9. Davidson, T., Warmlesley, D., Macy, M., Weber, I.: Automated hate speech detection and the problem of offensive language. In: *Proceedings of the International AAAI Conference on Web and Social Media vol. 11 No. 1* (2017)
 10. Dineva, D., Breitsohl, J.: Managing trolling in online communities: an organizational perspective. *Internet Res.* (2021). <https://doi.org/10.1108/INTR-08-2020-0462>
 11. Dineva, D., Breitsohl, J., Garrod, B., Megicks, P.: Consumer responses to conflict-management strategies on non-profit social media fan pages. *J. Interact. Mark.* **52**, 118–136 (2020). <https://doi.org/10.1016/j.intmar.2020.05.002>
 12. Dineva, D.P., Breitsohl, J.C., Garrod, B.: Corporate conflict management on social media brand fan pages. *J. Market. Manag.*, pp.1–20 (2017)
 13. Ditch: The Annual Bullying Survey. <https://www.ditchthelabel.org/research-papers/the-annual-bullying-survey-2018/> (2018)
 14. ElSherief, M., Nilizadeh, S., Nguyen, D., Vigna, G., Belding, E.: Peer to peer hate: hate speech instigators and their targets. In: *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 12, No. 1 (2018)
 15. Ewing, M.T., Wagstaff, P.E., Powell, I.H.: Brand rivalry and community conflict. *J. Bus. Res.* **66**(1), 4–12 (2013)
 16. Foody, M., Samara, M., Carlbring, P.: A review of cyberbullying and suggestions for online psychological therapy. *Internet Interv.* **2**(3), 235–242 (2015)
 17. Hutto, C., Gilbert, E.: Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8 No. 1 (2014)
 18. Ilhan, B.E., Kübler, R.V., Pauwels, K.H.: Battle of the brand fans: Impact of brand attack and defense on social media. *J. Interact. Mark.* **43**, 33–51 (2018)
 19. Kowalski, R.M., Giumetti, G.W., Schroeder, A.N., Lattanner, M.R.: Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. *Psychol. Bull.* **140**(4), 1073–1137 (2014)
 20. Kowalski, R.M., Limber, S.P., Agatston, P.W.: *Cyberbullying: Bullying in the Digital Age*. Wiley (2012)
 21. Leeftang, P.S.H., Verhoef, P.C., Dahlström, P., Freundt, T.: Challenges and solutions for marketing in a digital era. *Eur. Manag. J.* **32**(1), 1–12 (2014)
 22. Notar, C., et al.: Cyberbullying: a review of the literature. *Univ. J. Educ. Res.* **1**(1), 1–9 (2013)
 23. Pang, B., Lee, L., Vaithyanathan, S.: ‘Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprints/0205070* (2002)
 24. PEW Research Centre: Online harassment. http://www.pewinternet.org/files/2018/10/PI_OnlineHarassment_102214_pdf1.pdf (2018)
 25. Rodkin, P.C., Espelage, D.L., Hanish, L.D.: A relational framework for understanding bullying: developmental antecedents and outcomes. *Am. Psychol.* **70**(4), 311–321 (2015)

26. Sibai, O., et al.: Social control in online communities of consumption: a framework for community management. *Psychol. market.* (2015)
27. Singh, A.: Emotion detection using machine learning. <https://medium.com/@ParallelDots/emotion-detection-using-machine-learning-706ddceaa> (2017)
28. Slonje, R., Smith, P.K., Frisén, A.: The nature of cyberbullying, and strategies for prevention. *Comput. Hum. Behav.* **29**(1), 26–32 (2013)
29. Stephen, A.T.: The role of digital and social media marketing in consumer behavior. *Curr. Opin. Psychol.* **10**, 17–21 (2016)
30. text2emotion: text2emotion. <https://pypi.org/project/text2emotion/> (2020). Last accessed 2021/09/01
31. Tokunaga, R.S.: Following you home from school: a critical review and synthesis of research on cyberbullying victimization. *Comput. Hum. Behav.* **26**(3), 277–287 (2010). Available at: <http://linkinghub.elsevier.com/retrieve/pii/S074756320900185X>
32. Underwood, R., Bond, E., Baer, R.: Building service brands via social identity: lessons from the sports marketplace. *J. Market.* **9**, 1–13 (2001). Available at: <http://elibrary.ru/item.asp?id=6051799>
33. Van Heerde, H.J., et al.: Reality check: infusing ecological value into academic marketing research. *J. Market.* **85**(2), 1–13 (2021). Available at: <https://doi.org/10.1177/0022242921992383>
34. Whittaker, E., Kowalski, R.M.: Cyberbullying via social media. *J. Sch. Violence* **14**(1), 11–29 (2015)
35. Yadav, M.S., Pavlou, P.A.: Marketing in computer-mediated environments: research synthesis and new directions. *J. Mark.* **78**(1), 20–40 (2014)

Consumer Perception of Fast-Moving Consumer Goods About the Continente Food Lab Brand



João Pedro Carvalho and Ricardo Mena

Abstract The food retail market is very competitive. All players race against time to gain a competitive advantage over others. It is in this competitive environment that Continente launches a brand of innovative food products, Continente Food Lab. This new brand enters the market as a wager of Continente and SONAE MC in food innovation, however, it encountered some challenges in that same market, the main one of these challenges being its identity and the way in which it is perceived by consumers. If Continente Food Lab sees itself as a brand of innovative food retail products, consumers tend to perceive Continente Food Lab as a brand of healthy and organic products, which can bring some problems, with unfulfilled expectations on the part of consumers who wrongly expected to be buying healthy or organic products and a consequent loss of trust on the part of the consumer toward the brand. The purpose of this case study is to understand how the Continente Food Lab brand can overcome this problem of brand myopia it burdens. More specifically, this case study aims to define the attributes of the brand triangle that lead fast-moving consumer goods clients to perceive Continente Food Lab as an innovative food retail brand.

Keywords Marketing · Brand · Brand myopia · Branding · Retail · Food retail

1 Introduction

Marketing management must constantly adapt to today's unpredictable reality. Communicating their vision and purpose clearly is a growing need in companies, as is communicating their promises purposefully and consistently. It is also important that these companies are agile and flexible, giving more and more relevance to experimentation, testing, and innovation. SONAE MC is committed to the creation of a new brand, Continente Food Lab, focused on this paradigm of everyday life. The proposal of this new brand is to bring the food of the future to the consumer's kitchens.

J. P. Carvalho · R. Mena (✉)

IPAM Porto, R. Manuel Pinto de Azevedo 748, 4100-320 Porto, Portugal

e-mail: ricardo.mena@ipam.pt

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
J. L. Reis et al. (eds.), *Marketing and Smart Technologies*, Smart Innovation, Systems and Technologies 279, https://doi.org/10.1007/978-981-16-9268-0_55

653

During its legal evolution, the brand started to be considered as the visual representation of products. It was used for these same products to distinguish themselves, meaning, a tangible identifier of a unique product. Over the years, the need arose to make the concept of a brand something wider and wider. Today, the brand serves to identify a set of product and service offerings [1]. The authors distinguish three forms of myopia from branding. The first form of myopia is to perceive the brand as a distinctive product label; the second is to associate the brand with a specific offer. This is the type of branding myopia that is most often referred to in mainstream marketing books; finally, assessing the brand by the difference perceived by the customer. This is the type of branding that predominates in the literature.

In the 2004 special issue of the *Journal of Retailing*, the view that the retailer's brand is a broader concept than simply the branding of a product was discussed. Ailawadi and Keller [2] considered retailers' brands to be more multi-sensory than product brands. They further argued that the retailer's image in consumers' minds is the basis for their brand equity. The brand allows consumers to identify preferences and lower research and comparison costs. For these reasons, the retailer's brand, and the context in which it finds itself allows for similar identification and recognition, which mirror the level of trust on the part of the consumer, encouraging repeated purchase [3].

To help brands that face this problem of brand myopia, the author's de Lencastre and Côte-Real [1] conceptualized a brand model based on three pillars: the pillar of identity (which includes the sign or the group of signs that identify the brand.); the marketing pillar (which includes the offer of products, within the company and the marketing actions that support the brand); and the response pillar (which includes markets, in the sense of the different part of the brand, and their different responses to the brand). Since the value proposition of *Continente Food Lab* is not being perceived by the consumer, there is a need to understand what can be changed so that this perception also changes and the customer can easily identify that this brand is a promoter of products that are innovative and disruptive, rather than healthy and organic. The objective of this paper is to verify which changes in the brand triangle lead fast-moving consumer goods consumers to perceive the *Continente Food Lab* brand as innovative.

Research Questions.

1. What are the characteristics of the pillar of the identity that led FMCG consumers to perceive the *Continente Food Lab* brand as innovative?
2. What are the characteristics of the pillar of the market that led FMCG consumers to perceive the *Continente Food Lab* brand as innovative?
3. What are the characteristics of the pillar of the response that led FMCG consumers to perceive the *Continente Food Lab* brand as innovative?

2 Marketing Myopia

Many companies were threatened by innovation because they did not know how to adapt to these new forms of operating in their market. For example, Hollywood almost perished with the appearance of television. The reason that the appearance of television became a problem for Hollywood was the fact that film producers did not realize they were in the business of producing media for consumers and stuck on being in the business of producing movies, thus not taking advantage of the potential of television for the development of their business [4]. It can then be said that the term marketing myopia applies to explain a limited or distorted view of marketing when it considers that the products are merely the applied technology and not the benefit that customers can relish with the purchase of said product.

2.1 *Label, Product, and Customer, the Triad of Branding Myopia*

Label-branding myopia happens when excessive focus is placed on conceptualizing a very creative name or a very well-designed logo. In extreme cases, the meaning of symbology or name is completely neglected. For legal reasons associated with trademark protection laws, companies tend to combine syllables without any connection to the business area, but that sounds good together. This practice leads to a more difficult identification of which business area in which a given company or brand operates, thus causing a distortion of reality, through the brand name or logo [1].

On the other hand, product-branding myopia occurs when the brand is seen, in a more restricted way, as an integral part of the product. Many companies miss opportunities to expand brand reach by sticking to the market for the product they sell, rather than diverging to other markets with the same brand. Thus, companies see the brand only as one of the variables of the marketing mix, the product [5], and not as a broader asset. The Volkswagen brand, when expanding the range of cars, started to produce higher-end cars, for a target with greater economic strength. Brand decision makers questioned whether the brand name would not be a disadvantage to entering this high-end market, as Volkswagen translates to “People’s car”. At this time, the components and measures of brand equity began to be studied in greater depth, to better understand the brand’s value, such as Aaker’s Linear Learning Model [6].

Finally, the type of myopia that dominates branding books these days, the customer branding myopia. This kind of branding myopia happens when there is a special attention in considering the brand from the perspective of the seller versus the customer. In these cases, it is disregarded that the brand’s activity is not restricted to satisfying the interests of the products or customers. The brand does serve for the multiplicity of exchanges and relationships between a company and its stakeholders [7]. This myopia is a consequence of using a brand as a marketing tool and not as a company asset.

3 Brand Model

The brand model considers the triadic concept of the symbol, integrating it into the different areas of branding. The brand model is based on three pillars: the pillar of the identity, which includes the sign or group of signs that identify the brand; the marketing pillar, which includes the offer of products, with the company and the marketing actions that support the brand; and the response pillar, which includes markets, in the sense of the different pertinent of the brand, and their different responses to the brand.

Identity. We can consider the brand, in its narrowest sense, as the term or name expressed in a graphic form. But there are more legal issues when it comes to the brand. The brand identity mix is the set of name, logo, slogan, label, packaging, sound, and character. In addition, these identity signs are also constituents of legal identity.

The first time someone mentioned the concept of identity mix, it was to refer to four aspects of a company's identity management that had little to do with its image, and they were the facilities, products, communication, and people [8].

The authors of Lencastre and Côte-Real [1] identified the three levels of the identity mix, which are: (1) The Nuclear Identity, which is the symbol that the brand uses to identify itself in the first instance, usually the name; (2) The Real Identity, represents the graphic expression of the nuclear identity that has been registered, is the names, spellings and their logos (for example, the red cross name and logo change in Muslim countries to "Red Crescent"). (3) Finally, the Augmented Identity, which are all the other symbols of the brand that can be protected by law, such as a slogan, a bottle with a differentiated silhouette, a protagonist/mascot, or a sound.

Identity analysis begins with the company's identity mix, which makes it possible to access how a certain brand presents itself in a more direct way that is, its central identity; the way it graphically expresses itself that is, its real identity; and the remaining associated symbols and how they connect to each other (increased identity). It is even possible to understand whether it is a single or plural corporate brand identity structure by discovering whether any of these symbols refer to specific products and/or markets.

Marketing. In this pillar, it must be considered who is the commander of the destiny of a certain company at any time of its existence. This means that the physical or legal owner of the trademark, its name, and other recognized signs must be identified. That said, one must consider how the company is organized and how it allocates its resources to each product in its portfolio. Finally, the group of actions that favor the sale of each of these products in each market.

Marketing action is usually defined by the classic "4Ps" of marketing. This model defines the marketing action as product, price, place, and promotion [9]. Marketing myopia was later corrected by Kotler [5], when he presented the product variable in three concentric circles. The inner circle is the "Nuclear Product", which represents the core benefit of the product or service. The next circle is the "Real Product" which is defined by tangible benefits such as (1) brand name, (2) packaging, (3) quality, (4)

style, (5) and technological features. The outer circle is the “Augmented Product” which corresponds to additional services that accompany the purchase of the product or service itself and, usually, with ramifications for other variables of the marketing mix, such as after-sales services, home delivery, installation, and warranty.

Marketing analysis establishes what the brand presents as its core product, or the main object of its activity, as this product (main object of the activity) extends in multiple exchange relations. In a deeper analysis, it is possible to find supply actions that support the marketing mix, differentiating a specific product as the brand object (actual product) from all the actions that support it (the augmented product).

Response. In any conversation between two parties, an interpreter and a receptor are expected. There are different audiences (customers, suppliers, employees, shareholders, etc.) that can listen to a company’s message and, each one of them, can react and respond to that message in a different way. For Lavidge and Steiner [10], the concept of response, at the level of a group, has multiple meanings, including reactions in which the marketing language is classified as cognitive, affective, and conative, these meanings are explored later by Hofsted [11] in his study of business cultures and subcultures. Consecutively, Keller [12] applied this concept to the market, distinguishing the concepts of perception, preferences, and behavior.

In an attempt to organize the pillar of the response based on the triadic perspective, de Lencastre and Côte-Real [1] propose the following division by levels: (1) Nuclear Response, which is the immediate response from an individual when exposed to the symbol of the brand (It can be called the “top-of-mind” of the brand association or brand positioning); (2) the Real Response is proposed as being a more structured response that an individual has to the brand. The discourse of a given person, at this level, includes rival brands for sharing some similar points among themselves and also how these brands differ from each other (qualitative study of brand associations or brand image); (3) Augmented Response is the last level of this triad and corresponds to all possible reactions that the individual may have because they are perceptible at the level of real response from people in the same audience (quantitative area of market studies, share of mind, share of esteem, share of market, and quantification of brand equity). On all three levels, it is possible to distinguish cognitive, affective, and behavioral reactions. Response analysis that is, the response mix, can indicate the first word those consumers think about when exposed to the brand (top-of-mind association), using samples of brands or target segments. The primary response category is the brand’s positioning in that market or segment (core response). Then, all other spontaneous associations, which are detected in different qualitative studies, and expressed graphically in percentage maps, are quantified. The objective is to obtain a brand image compared to competitors (actual response). Finally, the brand value (augmented response) is established through the quantification of cognitive, affective, and behavioral responses, resulting from this image, and gathered in brand evaluation models.

4 Case Study

The case study involves an enormous challenge, especially when the investigation is of an exploratory nature. This method of doing research is an excellent tool when you want to document and delineate a list of possibilities or when trying to explain how you got those same possibilities. A case study usually involves the researcher in the study of a single case (although a small number of cases can be addressed in the same study) and using an experimental design [13].

4.1 Data Collection and Analysis

One way of doing qualitative research is through logic models. Logic models consist of expressing in a theoretical way the casual relationship between the data collected and the empirical knowledge found with the theoretical relationships stipulated initially. In thus, it is proven that an intervention produces results, comparing the research results with the results of the interviewed answers [13].

This investigation will have as a tool to obtain theoretical data the literature review, and as a tool to obtain empirical knowledge the interview with semi-structured questions to professionals in marketing, branding, design, innovation, and food retail.

4.2 Interview

For conducting the interview, the guidelines outlined by Yin [13] were taken as example, referring to interviews, in the context of scientific research, as relevant sources of information for a case study.

The process of creating the interview guide began with the elaboration of three sets of questions. Those sets correspond to each of the pillars of the brand triangle: Identity, Marketing, and Response. Subsequently, questions were created to obtain information to help build each of those pillars in all its own dimensions: Nuclear, Real, and Augmented. Delineate the participants is one of the main activities of scientific research, as it helps to qualify the quality of the interventions that are registered, and their relevance to the topic being addressed in the research [14]. Interviewed participants were a multidisciplinary group of 12 professionals from various fields, such as marketing, innovation, design, and food retail. This group is further complemented by Generation Y innovative food consumers, who for Yusoff and Kian [15] are from a generation very familiar with technology, digital, and innovation, as they are natives in an era of digital transformation, and consumers of innovative retail products from Generation Z, who according to Housand [16], are even more familiar with issues of innovation and technology than Generation Y, as they are the first generation to be born in a 100% digital world.

5 Data Analysis and Discussion

It is in this chapter that the answers obtained through the questionnaire to the different respondents are debated, related, or comparing them with the knowledge obtained through the literature review.

5.1 *Definition of the Identity Mix*

Nuclear Identity. The notion of brand identity varies depending on the authors who are consulted. For Aaker [17], brand identity is a set of associations that reflect what the brand implies. On the other hand, Nadan [18], the image and brand identity reflect development of the brand identity.

To assess this point, the interviewees were asked the question “What are the names that you associate with a food innovation brand in food retail?” The answers obtained do not show a great variation in the names that are associated with an innovative brand in food retail. A brand of innovative products in food retail must be associated with words such as “innovation”, “foodie”, “food lovers”, “alternative food”, “environment”, and “health”. Therefore, the Central Identity of a brand of innovative food retail products is identified with the words described above.

Real Identity. For Keller [19], brand knowledge is the personal meaning of the brand that is stored in the consumer’s memory. This knowledge includes descriptive and evaluative information. As for Esch [20], knowledge about a brand is stored in the mind in the form of a schema, which contains visual and verbal information, such as feelings, cognitive processes, and experiences.

This time, to assess this point, were asked two questions to the interviewees: “What should the logo of an innovative food retail brand look like?” and “What colors should represent an innovative food retail brand?” The logo design of an innovative food retail brand should be more minimalist, have as few distractions as possible from the word that is the brand name. The font of the text should be straight and thin. One of the suggested fonts to use in the logo was the Sans Serif font. Arrows, fruits with a double meaning (an inverted pear, that looks like a light bulb), and plants are some of the objects that should be used in the logo.

As for the colors to be used in the logo of an innovative food retail brand, they must be orange, green, or blue. All letters that make up the brand name must be the same color and size. Attention should be given to the contrast between the colors of the brand name and the logo so that the colors of the brand name have a positive contrast to the logo.

Augmented Identity. Keller [19] states that the best slogans are those that can contribute to the brand in different ways, for example, a slogan that informs consumers what the brand is about and that influences what they think about the brand. Carlsberg’s slogan is a superb example of this. “Probably the best beer in the world”, (1) connects the brand with its market, the beer market. (2) “...the best...”

invokes an association of being number one in this market. (3) Uses the word “Probably...” to generate sympathy and empathy for the brand. This slogan manages to influence consumers in their perception of the Carlsberg’s brand, and at the same time is a promoter of the essence of the brand (4). Functions 1–3 are constructs of how the consumer can be influenced and how they think about the brand, function 4 is a brand equity promoter [21].

To assess this point, the interviewees were asked the question “What symbols should be associated with an innovative food retail brand?” A good slogan for Continnente Food Lab would be something like “The most innovative food products, now on your table”. It was also referred mascots, such as in the shape of Erlenmeyer’s Balloon, inverted pear, and an owl in a scientist’s coat. In addition, sustainable packaging in the shape of a pear, made of sustainable materials are intended for online purchases, and packaging with a traditional format (cloth or raffia) for the purchase of food products in bulk or for multiple uses.

5.2 *Definition of the Marketing Mix*

Nuclear Product. A brand’s core product is the main trait that consumers associate with that brand when they encounter it. Most of the time, consumers do not buy goods or services just for the product itself. The decision of purchase goes through an assessment of a set of attributes that influence the value of the product according to its utility that is, benefits [22]. Other authors agree with this view. For example, Slater and Narver [23] state in their study that the value of the product to a consumer is built when the benefits are greater than the long-term costs that the consumer expects to have with that same product.

To assess this point, the interviewees were asked the question “What adjectives do you associate with the products of an innovative food retail brand?” and the main attributes linked with a brand of innovative food products are health, well-being, innovation, environmentalist, sustainability, and alternative.

Real Product. As for the Real Product, which represents the product as one of the variables of the marketing mix Lancaster [24] says that a brand can have the greater depth the greater the variety of products that this brand offers to consumers. He also says that offering more options has an advantage when compared to a range with fewer options, for the simple reason that a greater variety of options can cater to a wider variety of tastes. Other authors, such as Iyengar and Lepper [25], say that a smaller number of products in the range is more positive than a range with many products. In their study, consumers who had to choose a jam from a range of 6 were more likely to buy than those who had to choose from a range of 24 jams.

To assess this point, the interviewees were asked the question “What should be the products of an innovative food retail brand?” The real product of an innovative food retail brand focuses on defining its product assortment. In this investigation, it became clear that the products must be substitutes for meat and hydrates, which can promote greater sustainability and respect for the environment, while being

healthier alternatives than the products they intend to substitute. In addition, the product assortment should also consist of beverages that are less caloric, especially alcoholic beverages, or that do not contain alcohol, such as non-alcoholic wines. Finally, the assortment of products from a brand of innovative products in food retail must be aware of food trends, so that it can always be at the cutting edge.

Augmented Product. Increased product refers to all other actions that support the “P” of the product in the marketing mix that is, it refers to the remaining “P’s” of the marketing mix. For Singh [26], the “P” of the price is defined by the amount of money a consumer is willing to pay to purchase a certain product. The “P” for product refers to the physical product or service that a customer is ready to pay to obtain. Typically, a product is referred to as something tangible and service as something intangible. Finally, the “P” for place, or distribution, includes distribution channels, storage infrastructure, mode of transport, and inventory management and control.

To assess this point, the interviewees were asked the question “What activities can improve the shopping experience of an innovative food retail brand?”

Most respondents agree that the experience of purchasing innovative food retail products should be transversal to all distribution channels. The target audience for these products must be the Millennial and Generation Z generations. In the store, there must be specific stands or places for this type of product. Communication must be done through influencer marketing or endorsement with public figures or renowned chefs. In addition, the products would gain from their proliferation in well-known restaurant chains, making innovative food menus. The price should be generally cheaper, or it should promote discounts on a regular basis, to attract the consumer.

5.3 Definition of the Response Mix

Nuclear Response. The Central answer seeks to understand the positioning of the brand. The positioning of a brand is designated by the act of a company defining its offer and image to occupy a distinct place in the consumer’s mind [27]. According to a study by Eryigit and Eryigit [28], successful brand positioning can lead to greater consumer loyalty, greater brand equity, less vulnerable customers, better financial performance, and less price-sensitive customers.

To assess this point, the interviewees were asked the question “What are the sensations that awaken in you when you think of an innovative food retail brand?” The positioning of a brand of innovative products in food retail must be innovative, trendy, and healthy.

Real Response. The brand association is directly related to the image that each consumer generates in his mind in relation to a certain brand [29]. In fact, all information that represents the brand association is intrinsically connected to the brand name in the consumer’s mind and this is reflected in the brand image [30].

To assess which compounds are at this point, the interviewees were asked the question “What factors can differentiate an innovative food retail brand from its

competitors?” a brand of innovative products in food retail should be seen as the first to have the most innovative products available, a rapid stock rotation of products, a constant presence on social networks, physical proximity to customers and focus on products of national origin.

Augmented Response. The Augmented Response is about investigating the brand equity. This has become an important topic in the business world. For Aaker [31], there are four bases of brand equity that are related to the consumer, being brand loyalty (brand loyalty), name recognition, perceived quality, and other associations to the brand (brand association).

To assess this point, the interviewees were asked the following questions “Who buys products from an innovative food retail brand is it?”, “How do I feel when I buy products from an innovative food retail brand?” and “What do you think should be the shopping experience in an innovative food retail brand?” Consumers of innovative food products are seen by other consumers as avant-garde, curious, aware, informed, and courageous. In turn, when consumers buy innovative food products, they feel healthy, informed, adventurous, modern, and irreverent. Finally, they consider that the shopping experience of a brand of innovative food retail products should be simple and consumer-oriented, with fast and innovative processes, with useful information for the purchase and that the place of purchase should be easily accessible and differentiated.

6 Conclusions and Future Work

6.1 Conclusions

This paper contributes to the community with theoretical knowledge about the construction of a brand of innovative products in food retail, a brand that is focused on the consumer of these same products. Through scientific research, it was studied how a brand of innovative food products can present itself in the market so that will not be misinterpreted by consumers who buy products in food retail. In a simplistic way, there are brands that want to be seen in a certain way, but customers do not perceive them in the same way, thus generating myopia that is often caused by the brand’s investment in branding.

6.2 Future Work

This investigation does not have a definitive and finished quality as of the end of this document. The objective of this investigation is to open doors for knowledge of a subject that is still very little explored, the construction of innovative food product brands.

The greater limitation of this research was related to the inability to apply various data collection techniques for different research objectives. If there were access to these resources, the findings of the investigation would be richer. The definition of brand identity can be explored in a more sensory way. To get a better understanding of the real effect of its factors on human precession. In identifying the marketing mix of an innovative brand in food retail, more comprehensive data collection and analysis techniques can be used, such as the use of a surveys, to obtain a more focused response to from the public.

References

1. de Lencastre P, Côte-Real, A.: One, two, three: a practical brand anatomy. *J Brand Manag.* **17**(6), 399–412 (2010)
2. Ailawadi, K.L., Keller, K.L.: Understanding retail branding: conceptual insights and research propositions. *J. Retail.* **80**(4), 331–342 (2004)
3. Burt, S., Davies, K.: From the retail brand to the retailer as a brand: themes and issues in retail. *Int. J. Retail Distrib. Manag.* **38**(11/12), 865–878 (2010)
4. Levitt, T.: Marketing myopia. *J. Libr. Adm.* **4**(4), 59–80 (1984)
5. Kotler, P.: *Marketing Management: Analysis.* Prentice Hall, Englewood Cliffs, NJ (1967)
6. Aaker, D.A.: A new method for evaluating stochastic models of brand choice. *J. Mark. Res.* **3**(7), 300–306 (1970)
7. Balmer, J.M.T., Gray, E.R.: Corporate brands: what are they? What of them? *Eur. J. Market.* **37**(7/8), 972–997 (2003)
8. Pilditch, J.: *Communication by Design: A Study in Corporate Identity.* McGraw-Hill, Londres (1970)
9. McCarthy, E.J.: *Basic Marketing: A Managerial Approach.* McGraw Hill, New York (1960)
10. Lavidge, R.J., Steiner, G.: A model for predictive measurements of advertising effectiveness. *J. Market.* **25**(6), 59–62 (1961)
11. Hofstede, G.: Identifying organizational subcultures: an empirical approach. *J. Manage. Stud.* **34**(6), 947–952 (1998)
12. Keller, K.: Conceptualizing, measuring, and managing customer-based brand equity. *J. Mark.* **57**(1), 1–22 (1993)
13. Yin, R.K.: Validity and generalization in future case study evaluations. *Evaluation* **19**(3), 321–332 (2013)
14. Onwuegbuzie, A.J., Collins, K.M.T.: A typology of mixed methods sampling designs in social science research. *Qual. Rep.* **12**(2), 281–316 (2007)
15. Yusoff, W.F., Kian, T.S.: Generation differences in work motivation: from developing country perspective. *Int. J. Econ., Manag. Soc. Sci.* **2**(4), 97–103 (2013)
16. Housand, A.: CENTENNIALS: the world is waiting! Parent. *High Potential* **5**(2), 6–9 (2016)
17. *High Potential*, vol. 5, n° 2, pp. 6–9, 2016.
18. Aaker, D.A.: Measuring brand equity across products and markets. *Calif. Manage. Rev.* **38**(3), 102–120 (1996)
19. Nandan, S.: An exploration of the brand identity–brand image linkage: a communications perspective. *J. Brand Manag.* **12**(4), 264–278 (2005)
20. Keller, K.L.: Brand synthesis: the multidimensionality of brand knowledge. *J. Consum. Res.* **29**(4), 595–600 (2003)
21. Esch, F.R.: Brand identity: the guiding star for successful brands. In: *Handbook on Brand and Experience Management*, pp. 58–73. Cheltenham, UK, Edward Elgar (2008)
22. Rosengren S., Dahlén, M.: Brand slogan matching in a cluttered environment. *J. Market.* **12**(4), 263–279 (2006)

23. Caruana, A., Money, A.H., Berthon, P.R.: Service quality and satisfaction—the moderating role of value. *Eur. J. Market.* **11**(12), 1338–1353 (2000)
24. Slater, S.F., Narver, J.C.: Intelligence generation and superior customer value. *J. Acad. Market. Sci.* **28**(1), 120–128 (2000)
25. Lancaster, K.: The economics of product variety: a survey. *Mark. Sci.* **9**(3), 189–206 (1990)
26. Iyengar, S.S., Lepper, M.R.: When choice is demotivating: can one desire too much of a good thing? *J. Pers. Soc.* **79**(6), 995–1006 (2000)
27. Singh, M.: Marketing mix of 4P'S for competitive advantage. *IOSR J. Bus. Manag.* **3**(6), 40–45 (2012)
28. Kotler, P., Keller, K.L.: *Marketing Management*. Prentice-Hall, Upper Saddle River, NJ (2006)
29. Eryigit C., Eryigit, M.: Understanding the effectiveness of positioning bases with regard to customer perceptions. *J. Glob. Market.* **27**(2), 85–93 (2014)
30. Emari, H., Jafari, A., Mogaddam, M.: The mediatory impact of brand loyalty and brand image on brand equity. *Afr. J. Bus. Manag.* **6**(17), 5692–5701 (2012)
31. Aaker, D.: *Managing Brand Equity: Capitalizing on the Value of a Brand Name*. Free Press, New York (1991)

Gender Role Stereotypes as an Ethical Resource in the Peruvian Advertising Discourse



Ninowska-Camila Carbajal-Obando , Melina Mezarina ,
and Eliana Gallardo-Echenique 

Abstract The most renowned brands worldwide incorporate initiatives in their business approach that seek equality between genders in society. Nonetheless, brands in their advertising campaigns, gender remains stereotyped in the roles men and women play, which can be considered harmful to society and therefore unethical. Such discrepancy appears even in the same intervention generated by advertising specialists who define and design the communication of these brands. Therefore, this study seeks to identify the arguments used by Peruvian advertising professionals to justify the use of gender role stereotypes within the advertising discourse from an ethical perspective. A qualitative approach was selected as the research method for this study. In-depth interviews with 16 professional advertisers from different areas in charge of producing this type of communication. The lack of knowledge about the actual dimension of the harmful gender stereotypes impeded a debate to leave without arguments to those who believe that the use of stereotypes is ethical. Professional advertisers have lost part of the necessary connection with ethics and find it hard to deal with these dilemmas because they do not identify the damage caused by gender role stereotypes in society.

Keywords Stereotypes · Gender roles · Ethics · Advertising · Peruvian brands

1 Introduction

Unilever, one of the world's largest advertisers, carried out internal research that analyzes 1000 ads from 25 different countries and found that 50% of these ads

N.-C. Carbajal-Obando · M. Mezarina · E. Gallardo-Echenique (✉)
School of Communications, Universidad Peruana de Ciencias Aplicadas, Prolongación Primavera
2390, Lima 15023, Peru
e-mail: eliana.gallardo@upc.pe

N.-C. Carbajal-Obando
e-mail: u201513167@upc.edu.pe

M. Mezarina
e-mail: melina.mezarina@upc.pe

contained stereotypical portrayals of women, 3% showed them in leadership positions, and 2% showed them as intelligent [1]. Gender role stereotypes refer to the idea that certain behaviors learned by a person must be performed based upon their sex [2]. The advertising industry must contribute to combating gender stereotypes and sexism with positive messages toward women which must be perceived as authentic [3]. It is not only a trend that transmits unrealistic and unattainable stereotypes. Brands such as Procter & Gamble have received complaints about stereotyping advertising, portraying women as solely responsible for the purchase of cleaning and personal hygiene products [4].

A stereotype has traditionally been conceptualized and assessed as over-generalized belief about a particular group or class of people [5]. Gender role stereotypes are defined as behaviors and actions associated with a particular gender that influence the beliefs and expectations of others [6, 7]. Gender stereotypes represent women concerned for others rather than themselves and portrayed them as docile, vulnerable, fragile, modest, powerless, dreamy, child-like, warm, selfless, submissive, and helpless [8, 9]. Even today, the use of gender role stereotypes in advertising has negative consequences, especially in women, such as a loss of self-esteem, dignity, and confidence [10, 11]. Also, women have restricted opportunities for personal and professional development; they believe they cannot be successful in science or math to the point that they lose interest in Science, Technology, Engineering, and Mathematics (STEM) careers and affect their performance [10, 12]. Women's work is relegated and economically less valued even if it is the same position [6].

Therefore, this must be approached from an ethical perspective, understanding that it is present in every human act and is based on human rights inherent to all human beings [13]. Ethics is defined as the moral principles and values that govern people's behavior. Especially in advertising, ethics are not correctly aligned with legal norms, and these are the ones that erroneously indicate what is moral and what is not [4]. More than a decade ago, Drumwrigth and Murphy [14] were the first to discuss how advertising agency personnel perceive, process, and think about ethical issues. The authors found that advertising professionals have *moral muteness* and *moral myopia* because they failed to be concerned about advertising's social consequences. They have the obligation of examining, monitoring, and revising the standards and practices of ethical advertising [14]. Moreover, advertising is recognized as one of the most ethically challenging aspects of marketing [15]. Consequently, the discussion of the use of gender role stereotypes must be analyzed from the ethical perspective of the publicity experts. Therefore, the main research question of this study is: How do publicists justify, from the perspective of ethics, the use of gender role stereotypes in the advertising discourse?

2 Methodology

The aim of this study is to analyze how professional advertisers justify the use of stereotypes of female gender roles in Peruvian advertising discourse from an ethical perspective. A qualitative approach was chosen as the research method for this study [16]. A purposive sampling strategy was used to identify junior and senior professional advertisers who had experience in various brand categories such as creative (4), accounts (3), planning (2), marketing (2), art production (2), and direction (3). This sampling technique consists to find an informant who is willing to provide information due to the knowledge or experience he possesses [17].

In order to gather data answering the research question, an interview guide was applied with eighteen open-ended questions [18]. The interview guide was divided into three sections: (a) the use of gender role stereotypes; (b) the ethical aspects of gender stereotypes in Peruvian advertising; and (c) the ethical aspects of gender role stereotypes in the spot “saving’s experts”. In total eighteen questions. The spot for *Plaza Vea* supermarket “saving’s experts” was featured in interviews: Female voiceover accompanied by exciting music: What is a saving’s expert? A woman who knows about prices? A series of questions are presented about what an expert in saving is, relating them to motherhood, to being a housewife and a worker, all within the home. A sample of 16 interviews was obtained. The interviews, with an average duration of 45 min, were conducted in September and October 2020.

The thematic analysis was used to analyze the interview data. Once the data was transcribed, it was then analyzed, categorized, and organized into themes and further sub-themes [19]. All participants gave their verbal consent to tape-record discussion and to take part in an interview [20, 21]. The participants received a summary participant information sheet were informed that the study was completely voluntary, and would not affect their jobs, in any way [20]. To safeguard confidentiality, all interviewers were identified using an alphanumeric participation code.

3 Results

3.1 *Advertising Professionals’ Perceptions Regarding the Use of Stereotypes of Female Gender Roles*

About the stereotypes, the interviewees recognize incorrect characteristics and roles that typecast women and men just because of gender, where its acceptance is defined by the education received. The models that prevail are those of the wise homemaker and mother who takes care of children and the husband. Unfortunately, those occupations and behaviors are mostly attributed to women, which is harmful to them because their self-esteem is weakened and reinforces a sexualized image of them [22]. Even within this sexualization, women are assigned a less professional role and more demure role, which ultimately impairs the cognitive and emotional development and

their actions and attitudes [23]. After all, they create a misconception of reality that causes frustration and perpetuates its existence generating a vicious cycle where an oppressed behavior is reinforced. These external pressures are linked to lowered self-esteem and self-rejection feelings. The problem with such stereotypes is that they can shape wrong attitudes and perceptions of the world, by affirming or exaggerating such beliefs [24] or even can have a greater impact, because in addition to creating and perpetuating them, it harms gender equality and society, because it reinforces the gender focus and goes against equality [25]. Inside advertising, gender role stereotypes on women persist and tend to portray them as fragile, fearful, emotional, and dependent; but always showing grace, beauty, and sensuality. Interviewees did not believe that people were affected, or only were affected by causing angry in people against stereotypes. Furthermore, they attributed the impact to the business field and not the social field. On one side stereotypes enable people to identify themselves with a brand and establish a level of communication with a certain type of people; and on the other hand, it has to be careful enough to avoid damaging the brand's image and causing financial loss. The advertisers concentrate on depicting them in domestic roles such as housewives and beauty, recreation or ornamental models, and rarely in professional roles or a position outside the household or among other women.

In my experience, it has not affected them negatively, it worked well, but the advertising sought deception by making them believe that if they used a product they would be better moms, women, or wives. I do not think, it can affect on a psychological level. I do not know, what I do think is that if we had continued talking in that way, the effect would have been negative, because in some way, we are being responsible by telling women that they cannot be another thing. I have a daughter and I would like the message of a soap or detergent brand to reach her partner as well. (P8, M44)

I try to be as flexible as possible because above gender issues, there is the value of the person. (P10, W44)

A minority of interviewees claim their absolute opposition to any justification of the use of stereotypes because they consider that they do not contribute to society, create division, and take roots in the society. The motivation to incorporate gender stereotypes in advertising revolves around the consumer's identification with a certain brand, the fear of reduced sales, and the absence of studies to provide arguments to persuade them of the contrary. In line with Windels [24], all stereotypes are negative because they increase those pre-cognitions assigned to gender. These elements are used to reduce the message and the time it takes the receptor to sink in the information. Eisend [26] adds that the gender roles in advertising also affect those who do not consume the brand.

They are aimed at a profile that needs this type of advertising to understand it in a more didactic way. (PE5, W51)

Stereotypes are not justified. I do not feel that they contribute. I feel that they divide and somehow take root. I think that brands are very afraid of losing customers, and they betray themselves by not losing what they have built, by not losing sales. (PE9, M43)

3.2 *Ethical Arguments Regarding the Use of Gender Role Stereotypes in Peruvian Advertising Discourse*

The participants consider that ethics (a) depends on a personal judgment about what is morally correct; (b) consists of discerning between good and bad; (c) and seeks to be consistent with the values to protect the truth, rights, and duties to live in harmony in society. The majority of interviewees accepted that in some instances they had to struggle with it at work; today, they do not think about it because its limits are a blur, so they define them from their perspective. Advertisers outline a defense argument: the boundaries between right and wrong are a blur; that is why assessment criteria should be established and supported. Therefore, any ethical assessment becomes subjective. Something that could be perceived as ethical at the beginning could end up having the opposite effect. In line with Schauster [27], advertisers do not consider ethics in their actions, showing problems to conceptualize it; because it revolves around the legal.

According to Schauster and Neill [28], philosophical ethics considers that rational actions must be carried out from making decisions based on the understanding of what is right and what is wrong with it, and if it is morally justifiable or not. To Fine and Rush [29], ethics seek to provide humanity with “a good life” relying on the “what ought”. Therefore, they support the evaluations of suitability for gender-typed roles in advertising, arguing that the ethical value of gender equality should be a widely accepted ethical principle [29]. From the philosophical perspective, the concept of ethics establishes a set of standards on behavior that are considered “fair” or “right”. From there and by extension, emerges advertising ethics which focuses on what is “right” or “wrong”, even going beyond the legal field to qualify the actions and decisions as “good” or “bad” [15, 30]. That is because laws are the reflection of ethical judgment, so the advertising laws are dominated by advertising ethics and reaffirms that being legal does not mean that it is ethical [31].

Ethics. How difficult! (...) It is the behavior that you have in public and in private which aligns with values accepted as morally correct in the society, where predominates the truth and the common good. (P6, W56)

I have not seen ethics in a lot of years. I remember it from college. I believe that something ethical would be morally and legally correct, depending on the company principles where I work and that I have. (P13, W32)

It is complicated to define, you have your personal ethics and advertising ethics. There are certain moments in which you see and understand that perhaps advertising strategy and creativity are using exaggeration as a resource (...) So at what point do you say this is not ethically correct or socially it is not right? And you sin because you are a publicist, but you are also a person. It is complicated when you are already working, it is diffuse. (P3, M42)

About the female role gender stereotypes as ethical resources, a group considered them unethical because they create one expectation of reality, perpetuate the stereotypes and inequality, and do not allow to improve the ways advertisers communicate. Instead, another group considered that they are ethical as long as there are people who consciously accept the use of such stereotypes. These findings align with Tuncay

and Coleman [18], who found out that ethics is absent for advertisers as they put the responsibility of defining right and wrong on regulatory institutions. Advertisers felt they were just being creative or giving their customers what they want. Currently, the advertising world is ignoring the ethical principles that prevent them from perceiving ethical dilemmas or that they justify this practice [28]. According to Leter [30], practitioners rarely discussed ethical dilemmas even though they encounter ethical problems in their daily work; they prefer to ignore or minimize them. Therefore, ethics has to advocate not only for the consumers but also for the vulnerable persons in society [18]. I think it is ethical in the sense that it does not a lie when it [the spot] says that a woman is responsible for cleaning a home (...) but the morality is added here, it is correct? I think not, because we must be aware of what we are doing a prejudice and we can avoid it. For me, it is not moral, but ethical, because it is not a lie. (P11, M48).

3.3 Ethical Consideration about the Gender Role Stereotypes Present in Plaza Vea's Spot

Regarding the spot "saving's experts" of Plaza Vea (2018), the majority found it ethical because it empowered and portrayed women in a role in which they fought for equality; no male character was necessary for the campaign. In other words, the spot reflected reality. Interviewees questioned how saving money is attributed to the housewife, where her sacrifice as a devoted and exhausted mother who neglected herself is glorified. They conceded that probably the intention was good, but they missed the fact that when airing the advertising they reinforce negative stereotypes. In this case, gender role stereotypes excessively pigeonhole women that do not reflect the diversity of their lives [3]. Ethics denies any justification of the role gender stereotypes used in advertising. Ethics intends that "the fair" and "the right" prevail [13], which may harm gender equality and harm society [14]. Likewise, advertising is currently trying to use messages of empowerment and gender equality, even though sometimes it has a manipulative effect that apparently gives women a voice but actually reinforces the reproduction of stereotypes [8, 32].

I think that in that spot it uses a stereotype positively. Ethically, they are putting a stereotype in value. They are giving a woman strength and a voice. (...) Why does a woman have to be a saving's experts? Why cannot be a man? Why cannot I go shopping? I think that in the case of Plaza Vea perhaps it was an exaggerated approach, perhaps due to the search for an organic advertising impact. (E7, M34)

Yes, it is ethical because we are selling, this commercial goes to a specific target audience. But if you notice the self-sacrificing mother, in scenes she is very disheveled, she did not sleep (...) We could reduce it a bit. (E14, W29)

Likewise, the participants use often gender stereotypes despite being in disagreement or not sharing values with the brands they had to assess the risk of losing the account. One group highlighted that advertising should create and promote

campaigns to found alternative ways of communicating and fulfilling their responsibility with society. Meanwhile, another group justified the use of gender stereotypes in the advertising industry; because a large sector of the population felt identified with them, and the advertising had to reflect them to be effective. It is important to remember that stereotypical advertising perpetuates mistaken beliefs about female roles, and their repetition creates a perception that they are the only roles that women can play [7]. Likewise, these roles are mutually exclusive in a way that they are opposites [33].

It seems ethical to me because of how well managed it is by making the comparison between saving prices and saving in the struggle to achieve equality. I think a man must have appeared in this spot anyway. (E2, W37)

They practically assume that this woman lives in the house and is in charge of everything that happens there. I think that this intention to find common ground between saving with other aspects of her life seems somewhat forced to me and simplistic resource. Because under the pretext that the mother solves everything, stereotypes arise. (E12, M49)

4 Conclusions

The lack of knowledge about the actual dimension of the harmful gender stereotypes impeded a debate to leave without arguments to those who believe that the use of stereotypes is ethical. Professional advertisers have lost part of the necessary connection with ethics and find it hard to deal with these dilemmas because they do not identify the damage caused by gender role stereotypes in society. The interviewed justify the use of gender role stereotypes in advertising discourse as ethical because (a) the ethical limits are not established; (b) they have not received evidence of the harm that could cause; and (c) the ethical practices are often circumvented in daily practice.

Findings neither support a solid idea about how stereotypes can have harmful consequences to the brand's profits and image. It is necessary to prove what Lefter [30] claimed advertising has an economic and social impact but not individual. Findings evidence the need for an intervention managed not from an entrepreneurial field but academic to change managerial practices based on gender stereotypes. It is necessary to provide advertising and communications students with all necessary concepts and tools to reach an ethically acceptable position in which their actions are considered justifiable and sound. The relation between education and ethical dilemmas is evident. The universities must provide the necessary knowledge and skills to the new advertising professionals so they will be capable of critical thinking about possible moral and ethical dilemmas. The current absence of knowledge about the effects of sensitive ethical issues advocates the need to provide guidelines to navigate the ethics of personal versus professional behavior on practitioners in advertising.

This study has two limitations. First, the COVID-19 pandemic restrictions prevented the conducting of the traditional face-to-face interviews, so it was not possible to appreciate the non-verbal language of the interviewees who preferred the

camera turned off. Second, given that the data was collected from a small number of Peruvian professional advertisers, the findings are not generalizable in a different social context.

Taking these limitations into consideration, future research should study: (a) the level of stereotype threat, not only among the target consumers but also in society at large; (b) the impact on the current and future income trends and assets of the organizations that sometimes promote or defend the use of stereotypes; and (c) the final effects on the image of announcers and advertisers.

References

1. Sweney, M.: Unilever vows to drop sexist stereotypes from its ads (2016). <https://www.theguardian.com/media/2016/jun/22/unilever-sexist-stereotypes-ads-sunsilk-dove-lynx>
2. Cho, S., Jang, S.J.: Do gender role stereotypes and patriarchal culture affect nursing students' major satisfaction? *Int. J. Environ. Res. Public Health* **18**, 1–9 (2021). <https://doi.org/10.3390/ijerph18052607>
3. McCartan, A., McMahan, F.: Gender and advertising. *Int. Encycl. Gender Media Commun.* 1–17 (2020). <https://doi.org/10.1002/9781119429128.iegmc068>
4. López Jiménez, D., Dittmar, E.C., Vargas Portillo, J.P.: Self-regulation of sexist digital advertising: from ethics to law. *J. Bus. Ethics* **171**, 709–718 (2021). <https://doi.org/10.1007/s10551-020-04471-y>
5. Neuberg, S.L., Williams, K.E.G., Sng, O., Pick, C.M., Neel, R., Krems, J.A., Pirlott, A.G.: Toward capturing the functional and nuanced nature of social stereotypes: an affordance management approach. In: Gawronski, B.B.T.-A., in E.S.P. (ed.) *Advances in Experimental Social Psychology*, pp. 245–304. Academic Press, London, UK (2020). <https://doi.org/10.1016/bs.aesp.2020.04.004>
6. Clarke, H.M.: Gender stereotypes and gender-typed work BT—handbook of labor, human resources and population economics. In: Zimmermann, K.F. (ed.), *Handbook of Labor, Human Resources and Population Economics*, pp. 1–23. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-319-57365-6_21-1
7. Shinoda, L.M., Veludo-de-Oliveira, T., Pereira, I.: Beyond gender stereotypes: the missing women in print advertising. *Int. J. Advert.* **40**, 629–656 (2021). <https://doi.org/10.1080/02650487.2020.1820206>
8. Rodríguez, M., Gutiérrez, M.: Femvertising: female empowering strategies in recent Spanish commercials. *Investi. Fem.* **8**, 337–351 (2017). <https://doi.org/10.5209/INFE.54867>
9. Thompson-Whiteside, H., Turnbull, S., Howe-Walsh, L.: Advertising: should creative women be expected to 'fake it'? *J. Mark. Manag.* **37**, 294–319 (2021). <https://doi.org/10.1080/0267257X.2019.1707704>
10. Khalil, A., Dhanesh, G.S.: Gender stereotypes in television advertising in the Middle East: time for marketers and advertisers to step up. *Bus. Horiz.* **63**, 671–679 (2020). <https://doi.org/10.1016/j.bushor.2020.05.004>
11. Eisend, M.: A meta-analysis of gender roles in advertising. *J. Acad. Mark. Sci.* **38**, 418–440 (2010). <https://doi.org/10.1007/s11747-009-0181-x>
12. Steegh, A., Höffler, T., Höft, L., Parchmann, I.: First steps toward gender equity in the chemistry Olympiad: understanding the role of implicit gender-science stereotypes. *J. Res. Sci. Teach.* **58**, 40–68 (2021). <https://doi.org/10.1002/tea.21645>
13. De Melo Gomes, J.M.: *Compêndio Normativo da Actividade de Design* (2016). <https://doi.org/10.4995/Thesis/10251/61457>

14. Drumwright, M.E., Murphy, P.E.: How advertising practitioners view ethics: moral muteness, moral Myopia, and moral imagination. *J. Advert.* **33**, 7–24 (2004). <https://doi.org/10.1080/00913367.2004.10639158>
15. Shabbir, H.A., Maalouf, H., Griessmaier, M., Colmekcioglu, N., Akhtar, P.: Exploring perceptions of advertising ethics: an informant-derived approach. *J. Bus. Ethics* **159**, 727–744 (2019). <https://doi.org/10.1007/s10551-018-3784-7>
16. Creswell, J., Poth, C.: *Qualitative Inquiry Research Design*. SAGE Publications Inc., Thousand Oaks, CA (2018)
17. Etikan, I., Musa, S.A., Alkassim, R.S.: Comparison of convenience sampling and purposive sampling. *Am. J. Theor. Appl. Stat.* **5**, 1–4 (2016). <https://doi.org/10.11648/j.ajtas.20160501.11>
18. Tuncay Zayer, L., Coleman, C.A.: Advertising professionals' perceptions of the impact of gender portrayals on men and women: a question of ethics? *J. Advert.* **44**, 1–12 (2015). <https://doi.org/10.1080/00913367.2014.975878>
19. Clarke, V., Braun, V.: *Thematic Analysis* (2014). https://doi.org/10.1007/978-94-007-0753-5_3470
20. Rapley, T.: *Los análisis de la conversación, del discurso y de documentos en Investigación Cualitativa*. Ediciones Morata, S.L., Madrid, España (2014)
21. Deakin, H., Wakefield, K.: Skype interviewing: reflections of two PhD researchers. *Qual. Res.* **14**, 603–616 (2014). <https://doi.org/10.1177/1468794113488126>
22. Åkestam, N.: Caring for her: the influence of presumed influence on female consumers' attitudes towards advertising featuring gender-stereotyped portrayals. *Int. J. Advert.* **37**, 871–892 (2018). <https://doi.org/10.1080/02650487.2017.1384198>
23. Grau, S.L., Zotos, Y.C.: Gender stereotypes in advertising: a review of current research. *Int. J. Advert.* **35**, 761–770 (2016). <https://doi.org/10.1080/02650487.2016.1203556>
24. Windels, K.: Stereotypical or just typical: how do US practitioners view the role and function of gender stereotypes in advertisements? *Int. J. Advert.* **35**, 864–887 (2016). <https://doi.org/10.1080/02650487.2016.1160855>
25. Matthes, J., Prieler, M., Adam, K.: Gender-role portrayals in television advertising across the globe. *Sex Roles* **75**, 314–327 (2016). <https://doi.org/10.1007/s11199-016-0617-y>
26. Eisend, M.: Gender roles. *J. Advert.* **48**, 72–80 (2019). <https://doi.org/10.1080/00913367.2019.1566103>
27. Schauster, E.: Ethics versus survival: the relationship between advertising ethics and new business challenges. *J. Curr. Issues Res. Advert.* **40**, 90–104 (2019). <https://doi.org/10.1080/10641734.2018.1500320>
28. Schauster, E., Neill, M.: Have the ethics changed? An examination of ethics in advertising and public relations agencies. *J. Media Ethics* **32**, 45–60 (2017). <https://doi.org/10.1080/23736992.2016.1258993>
29. Fine, C., Rush, E.: “‘Why Does all the Girls have to Buy Pink Stuff?’” The ethics and science of the gendered toy marketing debate. *J. Bus. Ethics* 769–784 (2016). <https://doi.org/10.1007/s10551-016-3080-3>
30. Lefter, I.: Ethics and advertising—female targeted advertisement and the ethical concerns of practitioners and scholars. *J. Ethics Soc. Stud.* **2**, 33–42 (2019)
31. Förster, K., Brantner, C.: Masking the offense? An ethical view on humor in advertising. *J. Media Ethics* **31**, 146–161 (2016). <https://doi.org/10.1080/23736992.2016.1188013>
32. Varghese, N., Kumar, N.: Feminism in advertising: irony or revolution? A critical review of femvertising. *Fem. Media Stud.* 1–19 (2020). <https://doi.org/10.1080/14680777.2020.1825510>
33. Tabassum, N., Nayak, B.S.: Gender stereotypes and their impact on women's career progressions from a managerial perspective. *IIM Kozhikode Soc. Manag. Rev.* **10**, 192–208 (2021). <https://doi.org/10.1177/2277975220975513>

Consumer Involvement in the Co-creation of a Disruptive Offer: The Case Study of *Continente Labs*



Clésia Varandas and Mafalda Nogueira

Abstract During the last decade, the development of new technologies has revolutionized the operational models of the food retail sector. Despite the endless opportunities created by exponential technological growth, it has been increasingly challenging for companies to provide exceptional customer experiences. In this highly competitive context, retailers have been investing on shopping experiences that integrate technological environments to enhance and personalize the customer journey and also to optimize their operations. In addition, customers are very aware of their buying habits and, more than a simple purchase, they want to connect with brands. Thus, this research focuses on the consumer experience in the co-creation of a new and disruptive Sonae MC brand—*Continente Labs*—in which an analysis of the consumer experience in the online community was carried out. Through a qualitative exploratory methodology, we used netnography to analyze the participation and dynamics of the consumer in the *Comunidade Labs* platform. With regard to the customer cooperation platform at *Continente Labs*—*Comunidade Labs*—this study demonstrated customers increasing willingness to participate in co-creation processes, in the context of the firm’s strategy to attract customers. In this sense, through the study of the user experience in the online community, it was possible to explore and understand the experience of the collaborative consumer in a disruptive retail environment, which is developed through engaging, and personalized, experiences, allowing to connect customers and the brand, in order to satisfy their wishes.

Keywords *Continente Labs* · *Comunidade Labs* · Co-creation processes · Technological solutions · Customers’ collaboration

C. Varandas · M. Nogueira (✉)
IPAM, Porto, Portugal
e-mail: mafalda.nogueira@universidadeuropeia.pt

1 Introduction

Food retail is one of the most dynamic economic sectors, having witnessed, over the years, a significant change in consumer behavior and experience patterns, on one hand [1], and a constant adaptation from companies according to the needs and expectations of customers [2].

This new era, considered exponentially disruptive, is revolutionizing the relationship of retailers with consumers. In turn, customers, in their moments of consumption, more than a merely transactional relationship, look for a continuous and personalized journey. According to the “*Consumer Connectivity Insights*” report, today’s consumers are increasingly receptive to more connected experiences. This study indicates that 69% of consumers believe that a disconnected consumer experience would make them switch service providers [3].

Thus, this investigation results from a research project carried out at Sonae MC, with the aim of analyzing the *Continente Labs* brand, which adds a space for direct communication with the end customer called *Comunidade Labs*, which, through an interactive online platform, aims to integrate consumers, as an active part, in the dynamics of the process of co-creation and co-development of new products, applications or store spaces. In this sense, the main goal of this paper is to explore and analyze this co-creation process through the consumer’s experience and participation in the online community.

2 Literature Review

2.1 Consumer Participation in Co-creation Processes

In the modern, highly interconnected world, the role of consumers has changed substantially due to their active collaboration with companies in product and service innovation [4]. In addition, it is possible to confirm that consumers are interested in having a more active role in the consumption process and in the interaction with their own organizations [5]. It can be said that, currently, customers are not part of a “passive audience,” but of “active co-producers,” because they show interest in the co-creation process through their interactions with products, brands, and companies [6].

2.2 Customer Focus

Despite the complexity of the topic in question, several trend studies argue that the retail market should listen to, and integrate, the customer in the construction of new solutions and respond, accordingly, in a customer-centric logic, considering that 80% of the retail aims to create mechanisms of direct interaction with customers in real time [7].

As suggested by Mukherjee and Banerjee [8], this customer-centric orientation provides the appropriate framework to leverage retail innovation. Furthermore, developing emotional bonds with consumers can be an excellent strategy to drive innovation in the shopping journey [8].

2.3 Online Communities

Social networks or other online channels of interaction with the customer contribute to support innovation, as they facilitate the approximation of entities [9]. From this point of view, companies are increasingly betting on technological platforms for online communities, in order to engage with customers and generate new ideas, in a logic of value co-creation and co-development between the company/brand and their customers/participants [10].

Given this logic, several retailers, instead of presenting products and services in their final conceptualization stage, opt for open innovation dynamics and offer ways to create new ideas and mechanisms with their customers [11]. To allow for this type of co-creation dynamics, interactive platforms are developed, such as online communities, around activities related to retail [12].

Authors argue that online communities are social collectives marked by a shared culture, where the brand identity is collectively represented, negotiated, and co-created through the meanings, rituals, and consumption practices of the members [13]. In turn, community members are brand admirers who are willing to express their perceptions [14].

Muniz Jr. and O'Guinn [15] identified four crucial types of relationships in online communities emerging through interactive experiences between: the customer and the brand; the customer and the company; the customer and the brand product/service in use; and among other customers [13]. In this sense, managers must have an understanding of the sociocultural influences of collaborative dynamics and influence co-creative activities [16].

In practice, companies such as *Dell*, *Starbucks*, *Mercadona*, *Honeywell International* and *Maersk Line* have built online communities to successfully market products and services [17]. Although in the retail world, there are some successful examples of this type of online communities, there are also organizations that fail to take advantage of online innovation communities. This is due to the high rate of member turnover, which generates ineffective collaboration processes among participants [18]. Co-creation is often discussed in the context of an online community where members can play an active role, both as providers and beneficiaries [19].

2.4 Company–Customer Connection and Value Co-creation

Value co-creation is an abstract concept, and recent studies call for microfoundational research to understand how it is operationalized in business practices. However, in

the context of collaborative innovation communities, it represents the value that is co-created, leveraged by members, and through interactive technology platforms [20].

Although companies try to employ various stakeholder retention strategies, such as reward systems, competitive initiatives, sometimes it may not be enough to capture the interest of the online community [21]. Most empirical studies in this context consider customer engagement in collaborative innovation from an organizational perspective [22]. That is, it captures the benefit to the organization and does not explain the co-created value for the individual resulting from collaborating to obtain self-value [23]. However, it is essential for companies to analyze the factors that condition individuals to participate in this collaboration, in order to ensure the desired involvement in their online communities [24]. Thus, researchers recognize the importance of this study on the management of online communities and the individuals who integrate it to obtain a deeper understanding of the value co-creation activities carried out on online collaborative innovation platforms [25]. Here, the aim is to create a secure connection between the brand and its users [26].

Typically, community members are usually like-minded individuals who share hobbies and mutual interests in the brand/company. This involvement is usually marked by interests and reasons that are important to the members [27]. In this way, the ongoing relationship enhances collaborative innovation through the organic activities of the community, namely in the sharing of user experiences, in exposed questions related to the product/service, in the offer of problem solving, and contributing with new ideas [28].

Von Hippel [29] argues that innovation communities should be an open door for all individuals who wish to contribute to the development or improvement of a product or service, not necessarily for current customers, even because the inclusion of this type of participants leads to a broader spectrum of ideas [30].

Several studies claim that this type of value co-creation collaboration through online platforms can have very positive results for the companies' innovation process and, consequently, generate an increase in revenue [31]. This process requires the involvement of community members in a series of activities, such as presenting ideas, sharing knowledge, and providing feedback. In practice, this dynamic allows the participation in voting activities, such as selection of products to be marketed, review and suggestion of titles or descriptions, and exploration of the experimental base in a real environment with the product or service, such as commenting or identifying new sources of innovation [32].

Basically, community members become creators, designers, and intermediaries, given their interaction and learning between each other's shared experiences [24]. In more specific cases with a strong link in the experimentation aspect of solutions composed with a certain digital literacy, they can be called beta testers.

Consumer participation then became essential as a source of innovative ideas and brand value. However, although existing research identifies the different elements involved in consumer co-creation, there is still a need to better understand this mechanism. Accordingly, some of the biggest players in the market (*Walmart to Nike*) are embracing new idea accelerators in creating interactive cooperative innovation

labs to connect with consumers bilaterally in a digital setting [33]. This new level of integration and consumer acceptance enhances new opportunities for brands to conceptualize strategies that are more in line with customer needs [4].

3 Conceptual Model

Currently, the phenomenon of digitization is one of the most important transformations that retail, in general, is facing. It has not only drastically changed business opportunities, but also business models, purchasing processes, and forms of consumption. Basically, the diffusion of digital technologies transforms the way retailers present and provide new products and services to consumers as well as new forms of consumption associated with the use of these digital technologies. In this way, the proposed conceptual model in Fig. 1 makes it possible to logically systematize the dimensions of analysis suggested in this study. Thus, the model represents some of the dimensions inherent to consumer participation in value co-creation dynamics, with a focus on the customer, online communities, and connection between company and customer [4].

In turn, as shown in Fig. 1, it was considered pertinent to divide and intersect the conceptual model into two large dimensions, namely one for the digital environment associated with Community Labs, and another for the physical environment related to Contiente’s disruptive store Labs—as they are directly interconnected, designing a truly collaborative experience between physical and digital media. According to Pantano et al. [34], the union of the online and physical environment aims to offer an interactive, instantaneous, and more satisfying experience to users [4].

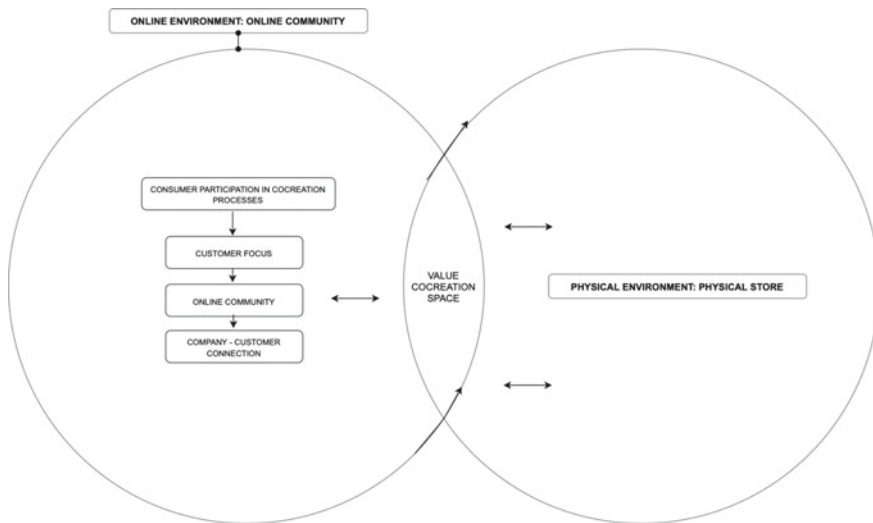


Fig. 1 Conceptual model

4 Methodological Decisions

4.1 Methodological Approach

As this is a complex topic, it adopts an exploratory qualitative method based on a case study of the *Continente Labs* brand in the company *Sonae MC*. Qualitative research is a holistic approach that involves discovery and can be defined as any type of research that does not present results through statistical procedures or other means of quantification. That is, instead of measuring the phenomenon of integration by numbers, we intend to use open-ended questions to explore possible perspectives [35]. Furthermore, it is described as a model that occurs in a natural environment and allows the researcher to develop a level of detail that is highly involved in real experiences [36]. Although the research is primarily inductive and supported by a theoretical framework, the data should guide the study, not a theory [37]. As such, the focus of this methodology is on deep understanding of words, opinions, and experiences, rather than numbers [38]. In turn, it is considered a subjective approach because it is based on the interpretation of phenomena in the context in which they live, that is, this means that the researcher must be directly involved with the situation under study and have the capacity to interpret to analyze and constructively criticize the observed/captured events [39].

4.2 Data Collection

Considering that the *Continente Labs* brand adds a space for direct communication with the end customer in an interactive online platform—the *Comunidade Labs*—which aims to integrate consumers as an active part in the dynamics of the co-creation and co-development process of new products, applications or store spaces, netnography was defined as the data collection method to analyze consumer participation in digital dynamics of value co-creation.

In fact, the emergence of the digital network organized into online communities, forums, and blogs has changed the way consumers communicate and has become a powerful communication channel for digital consumers [40]. Within these online environments, it is possible to observe the data generated and accessible by the user about desires, wishes, beliefs, and experiences of consumers in relation to products, services, and perceptions of the brand and the company [41].

Therefore, both research methods focusing on analyzing existing consumer dialogue and interactions in consumer communities are increasing. Based on the heritage of ethnographic research, netnography is a method that has evolved as an observational instrument that is applied in digital environments, with the term “netnography” being a neologism derived from the words “internet” and “ethnography” [41].

One of the main reasons for choosing this technique in response to the research problem is that consumers voluntarily reveal information, sensitive and unsolicited

details. Effectively, netnography has the reputation of being one of the most effective techniques for observing consumer behavior, in digital environments, understanding their tastes, desires, and factors influencing customers in their consumption decisions [42].

Thus, the need for a systematic process to identify, select, and analyze online consumer interactions emerged. From this perspective, the netnographic analysis was classified as analytical [43]. Thus, through the analytical netnographic analysis, two main moments of observation were considered, exposed in the following topics. In this sense, numerical data related to platform users, members participating in challenges exposed in the community, page views, member interactions through comments or likes, and the average duration of the session were observed, registered, and analyzed. This phase of collecting and analyzing netnographic data is essential to understand if the online community attracts participating members [4].

On the other hand, it is worth noting that given the high number of comments and likes that *Comunidade Labs* presents and the platform's inability to extract this content into a database, it was not possible to qualitatively analyze each of the corresponding interactions.

In any case, it is intended that the researcher takes a critical look at the analyzed content in order to find various information about the perceptions of consumers around the *Continente Labs* universe. In this way, online interactions are perceived as a valuable cultural reflection, producing deep human understanding [4].

4.3 Universe of Analysis and Participants

The universe of this analysis is composed of individuals who are involved with the *Continente Labs* brand. However, this universe is made up, for the most part, of *Continente Labs* brand customers or members participating in the online value co-creation platform of that brand—the *Comunidade Labs*. Through a segmentation strategy carried out by the researcher with the *Continente Labs* team, it was possible to segment the profile of the brand's associated customers, which are divided into three types of customers with different characteristics, as shown in the table in annex 1. It should be noted that the segmentation shown was defined under the exclusive criteria of the *Continente Labs* brand. However, the researcher used this segmentation to organize data collection.

4.4 Data Analysis

After the first phase of data collection, the information gathered was explored and organized using the content analysis technique. This technique involves organizing the information into categories according to the main topics and questions of the study. Here, the researcher must demonstrate the ability to identify pertinent information and separate it from those that are not pertinent [35].

Table 1 Continente Labs customer segmentation

	Analysis Periods	Analysis date
First moment	Considers the period from its creation to the last day defined for data extraction	04/01/2020 until 06/30/2021
Second Moment	It considers the implementation date of the first internal customer segmentation strategy until the last day defined for data extraction	01/25/2021 until 06/30/2021

According to Hsieh and Shannon [44], this procedure contains three main steps: (1) pre-analysis, (2) exploration of the material, and (3) treatment of results: interference and interpretation. In addition, at the beginning of the pre-analysis phase, all notes and transcripts were read to obtain an overview of the body and context of the data collected [44].

Then, the reading process was followed by a coding process, in the material exploration phase, in which each category was systematically labeled according to the relevance of the data and the theoretical framework of the literature review. Finally, in the data processing stage, the researcher critically analyzed the results, seeking possible conclusions to respond to the object under study.

5 Results

5.1 Descriptive Analysis

Following the established netnographic matrix, the analysis of *Continente Labs* platform data was first developed through the exploration and extraction of various numerical data related to the platform users, page views, member interactions through comments or likes, and the average duration of sessions in order to realize the attractiveness of this same platform.

As mentioned earlier from the analysis of the *Comunidade Labs* analysis, and represented in Table 1, the extraction of data from the analytical netnography considered two main moments of observation:

5.2 Content Analysis

At the first moment of analysis, *Comunidade Labs* presents 4923 visits to the site (see annex 2), taking into account that this algorithm is counted only once and does not double the value from the moment that identifies the user's IP address. In contrast, it identifies 14,486 page views, and this metric already counts the number of total views on the website; even when the user views the page more than once a day,

all website views are counted. Regarding the average duration of sessions, which represents the time visitors spend on the *Comunidade Labs* platform, it is 00:00:1:50, indicating a rejection rate of 78.47%, which reflects the percentage of users who abandoned the website without interacting after viewing the main page. As for the number of interactions, during this period, 1759 comments, 1086 likes, and 207 new topics were created.

Following the same order of analysis, in the second defined period, the *Continente Labs* platform presents 4419 total visits, 10,405 page views, 00:01:07 average session duration, 81.05% bounce rate (see Annex 3). With regard to the number of interactions, 1447 comments, 781 likes, and 73 topics created were identified. It is still possible to see that most users of the digital laboratory are concentrated in the Lisbon region, with 1563 users, and in Porto, with 966 users, in both moments of analysis.

Taking into account the two periods of analysis, both register the January 26 as the most affluent occasion in the Community of Labs, a consequence of the implementation of the first customer acquisition campaign in an exclusive newsletter and a banner in the “*Cartão Continente*” application, with 702 page views and 441 gross visits to the *Comunidade Labs* website.

In turn, it is possible to affirm that the increase in the activity of *Comunidade Labs* corresponds to the second period of analysis. The number of users increased by 4419 visits to the platform between moments of analysis, that is, from 504 (first moment) to a total of 4923 users (sum of the first moment with the second moment).

In the case of total views on the website, the situation is identical; it presents an increase of 10,405. Since in the first moment, it estimates 4081 views and with the records of the second moment, it defines a total of 14,486 views. Regarding the metrics that determine the number of interactions, in the first moment, there were only 312 comments, 305 likes, and 134 topics created, and the total of interactions added with the second moment resulted in 1759 comments, 1086 likes, and 207 topics created. As for the rejection rate, although in this case, it is not a positive situation because it corresponds to the number of sessions that left the site in relation to the total number of sessions performed; it also increased from 78.47 to 81.05%. Likewise, the average length of time that visitors spend on the website has decreased.

Since the *Comunidade Labs* is integrated in the *Continente Labs* website, it was considered relevant to check the behavior of users in this installment and, thus, check the number of events triggered for the respective co-creation platform. Thus, as shown in annexes 6 and 7, in the first moment, a total of 4699 clicks were registered, and 4156 clicks in the corresponding second moment, that is, there was an increase of 543 clicks between periods of analysis.

Since the traffic overview graph on the *Continente Labs* website highlights the greater flow of views on May 26, which results from the opening day of the *Continente Labs* physical store, a third moment of analysis was considered, from May 26 to June 30, 2021, in order to understand if the phenomena were correlated. Thus, as shown in annex 8, a total of 501 clicks were recorded from the *Continente Labs* website for the *Comunidade Labs*, bearing in mind that 109 clicks were triggered on the opening day of the *Continente Labs* physical store alone. However, although some actions

were signed on the opening day of *Continente Labs* store, the researcher believes that this event was not a crucial element to increase the activity of *Comunidade Labs*.

6 Discussion and Conclusions

6.1 Main Findings

Taking into account the two main analysis moments defined in the netnographic analysis of the *Continente Labs* value co-creation platform—*Comunidade Labs*, it was found that *Comunidade Labs* presents 4923 visits, 14,486 page views, 00:001:50 average session duration, and 78.47% bounce rate. Regarding the number of interactions, 1759 comments, 1086 likes, and 207 topics created were identified. Although the data presented in the bounce rate and average session duration metrics correspond to unambitious values, the increase in *Comunidade Labs* activity is marked from the second period of analysis when the customer acquisition strategy is implemented. In response to the initial question of this investigation—how the involvement of consumers contributes to the co-creation of the *Continente Labs* brand, through their experience in the online community—it is possible to say that companies have sought to respond to current challenges using innovation accelerators digital and experimental as digital laboratories, in order to generate new ideas with the collaboration of customers. The consumer experience in the online community is developed by the strong connection between customers and brands, in the entire universe of the value chain, in a logic of co-creation, co-production, and co-development. Thus, for the *Continente Labs* brand to thrive in a collaborative environment, it needs to create continuous experiences that involve the customer and boost their collaboration. Furthermore, it is essential that the brand builds a relationship of extreme trust and transparency with the customer, in which it shares direct and fluid communication.

6.2 Contributions of the Study

From a theoretical perspective, this research contributes to deepening the knowledge of the food retail sector, which invests in technological solutions through collaborative dynamics in the customer's purchase journey.

In fact, this research contributes to a differentiating analysis, as there are few concepts in food retail that encompass the laboratory component of co-creation. In fact, it has been verified that the market feels the need to collaborate with consumers themselves and some cases of brands combined with co-creation and co-development laboratories have been identified, but normally, this experimental component is integrated only in a single environment, either in the physical environment or in the online environment.

Moreover, the proposed model is pioneer because it reflects a very specific technological component in the academic literature and, consequently, recent in the market. Thus, this conceptual model can be considered unique because it has the ability to perceive consumers who were part of the experience as co-creators on a platform for this purpose.

From a practical point of view, and during the research project, this study contributed with several actions to *Continente Labs* through the implementation of engagement dynamics to maintain high levels of interest in the brand and participation in the *Comunidade Labs*. This study also collaborated in the definition of a strategy to attract customers to participate in *Comunidade Labs* in the implementation of an exclusive newsletter and banner in the application of the “*Continente Card*” aimed at customers with the characteristics mentioned in annex 2. In turn, support actions were also implemented for product managers that introduce experiences in the *Comunidade Labs* with proposals for newsletters for platform members and posts from social networks with the challenges in vogue. Furthermore, this research indicated several metrics capable of analyzing the interaction level of the *Comunidade Labs* platform.

Likewise, some recommendations to the brand are listed. Although the tool chosen for the construction and maintenance of the website is restrictive, with limited services at the company’s option, based on the agile methodology, in which user inputs are lateral in prioritizing and developing new features, this research recommends that user interface (UI) and user experience (UX) techniques be implemented in order to improve user experience, interest and interaction.

Likewise, it is recommended that the *Continente Labs* team bet on collaborators who are totally focused and dedicated to managing the activities of the *Comunidade Labs* because being a platform that communicates directly with the end customer, it needs to remain present and available in order to generate ongoing engagement. Otherwise, customers feel that their suggestions do not gain traction and automatically lose interest in co-creating and co-developing Continental dynamics. In view of this, it would be advantageous for the team to continue to invest frequently in social media tools to communicate brand news, both from the *Continente Labs* store and from *Comunidade Labs* itself.

It is noteworthy that all the suggested points arise from situations analyzed by the researcher in loco, with the *Continente Labs* team, mostly based on the data gathered when the technique and data collection established in the investigation was carried out through several publications by Community Labs users.

6.3 Limitations and Suggestions for Future Research

This investigation took into account the care for scientific and methodological rigor; however, there are some inherent limitations that will be exposed, accompanied by suggestions for future investigations.

Regarding the limitations, although this investigation meets the criteria of coverage, the number of participants is reduced due to the useful time to carry out the investigation, taking into account that it was only possible to collect most of this data during a short period of time.

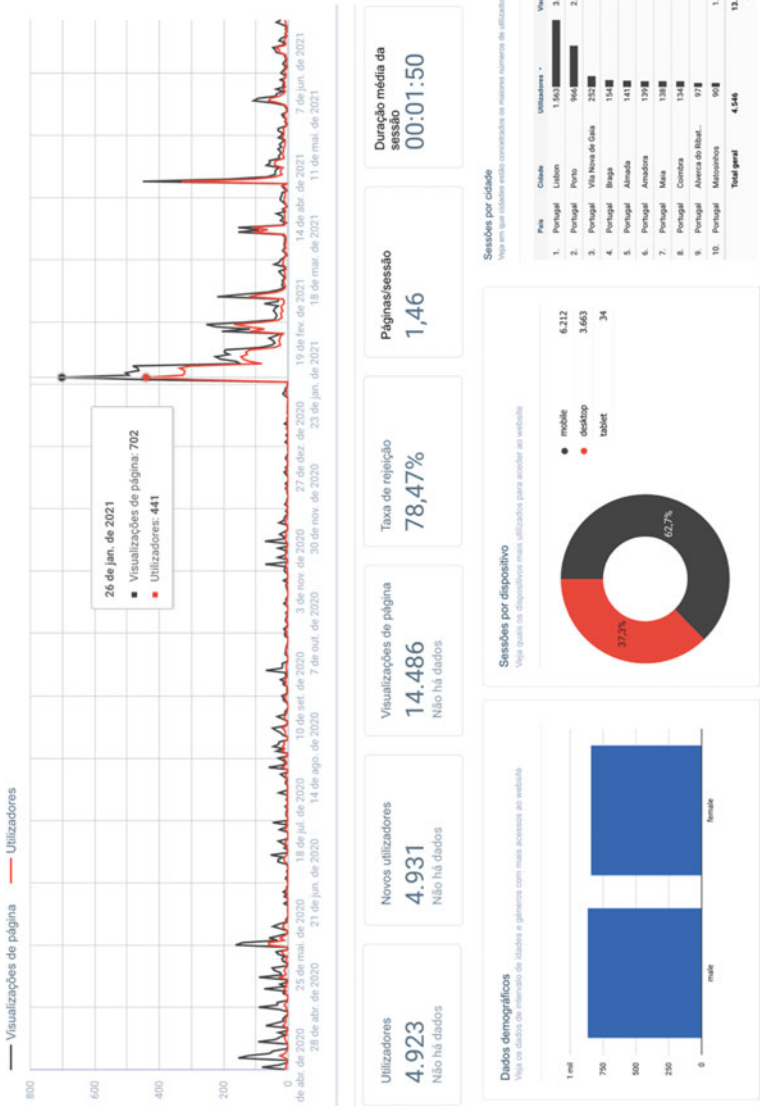
As suggestions for future research, taking into account that the topic addresses a little-explored trend, it would be pertinent to deepen the objective defined in this research using the quantitative methodology, in carrying out surveys of *Continente Labs* participants and customers, in order to assess consumer participation in value co-creation dynamics.

Attachment

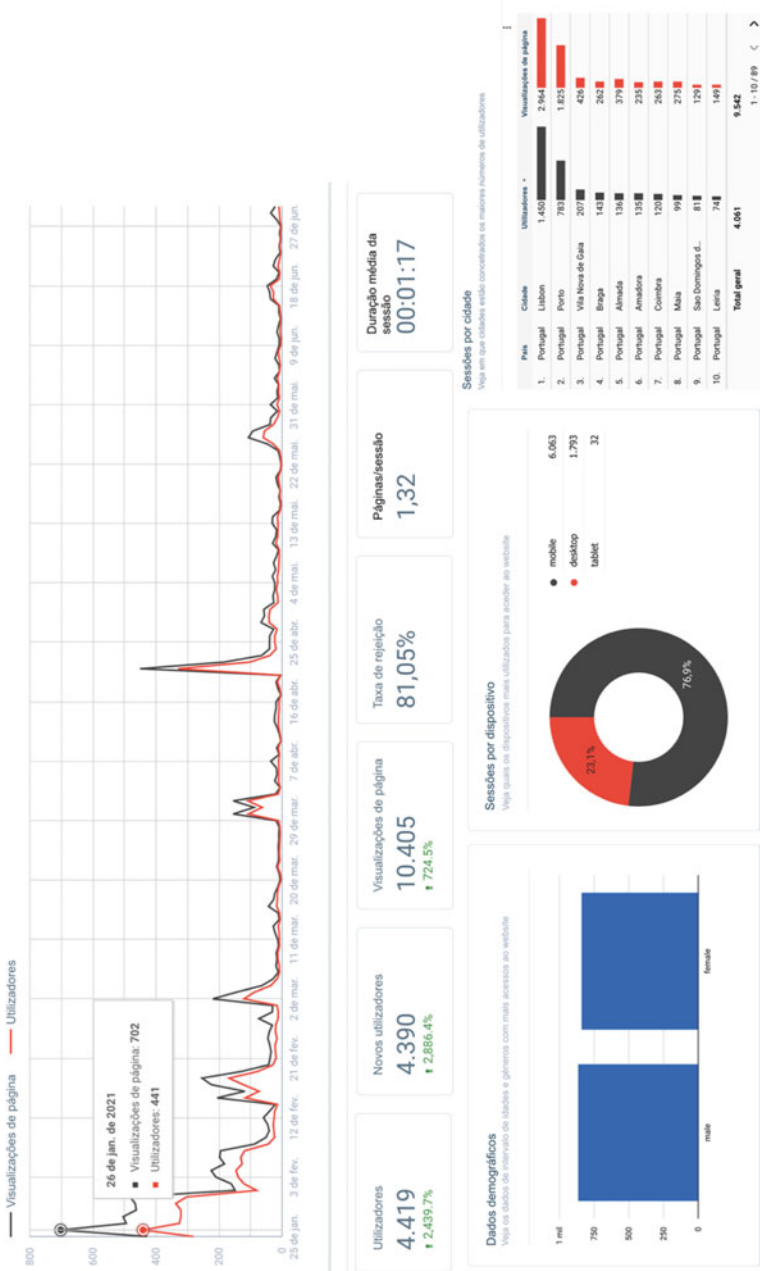
Annex 1: Continente Labs Customer Segmentation

Customer A	Demanding healthy and sophisticated urban people Loyal large and loyal medium customers High and medium share of wallet Customers who used the “ <i>Continente Card</i> ” application more than once in the last three months Ages between 30 and 45 years old Basket with Bio & Healthy products, Prozis, Fruits and Vegetables
Customer B	Demanding healthy and sophisticated urban people; Loyal large and loyal medium customers High and medium share of wallet Customers who used the “ <i>Continente Card</i> ” application more than once in the last three months Ages between 20 and 35 years old
Customer C	Early adopters/trendsetters customers

Annex 2: Comunidade Labs Dashboard (1st moment)



Annex 3: Comunidade Labs Dashboard (2nd moment)



References

1. Vrontis, D., Thrassou, A., Amirkhanpour, M.: B2C smart retailing: a consumer-focused value-based analysis of interactions and synergies. *Technol. Forecast. Soc. Change* **124**, 271–282 (2017). <https://doi.org/10.1016/j.techfore.2016.10.064>
2. Adapa, S., Fazal-e-Hasan, S.M., Makam, S.B., et al.: Examining the antecedents and consequences of perceived shopping value through smart retail technology. *J. Retail. Consum. Serv.* **52**, 101901 (2020). <https://doi.org/10.1016/j.jretconser.2019.101901>
3. Webb, A.: 2018 Emerging Tech Trends. *Future Today Inst.* 1–248 (2018)
4. Albors-Garrigos, J.: Barriers and enablers for innovation in the retail sector: co-innovating with the customer. A case study in grocery retailing. *J. Retail. Consum. Serv.* **55**, 102077 (2020). <https://doi.org/10.1016/j.jretconser.2020.102077>
5. Gupta, S., Leszkiewicz, A., Kumar, V., et al.: Digital analytics: modeling for insights and new methods. *J. Interact. Mark.* **51**, 26–43 (2020). <https://doi.org/10.1016/j.intmar.2020.04.003>
6. Blanco-Callejo, M.: Co-innovation at Mercadona: a radically different and unique innovation model in the retail sector (2019)
7. Hajli, N., Shanmugam, M., Papagiannidis, S., et al.: Branding co-creation with members of online brand communities. *J. Bus. Res.* **70**, 136–144 (2017). <https://doi.org/10.1016/j.jbusres.2016.08.026>
8. Mukherjee, K., Banerjee, N.: Social networking sites and customers' attitude towards advertisements. *J. Res. Interact. Mark.* **13**, 477–491 (2019). <https://doi.org/10.1108/JRIM-06-2018-0081>
9. Pantano, E., Priporas, C.V.: The effect of mobile retailing on consumers' purchasing experiences: a dynamic perspective. *Comput. Human Behav.* **61**, 548–555 (2016). <https://doi.org/10.1016/j.chb.2016.03.071>
10. Bosch-Sijtsema, P., Bosch, J.: User involvement throughout the innovation process in high-tech industries. *J. Prod. Innov. Manag.* **32**, 793–807 (2015). <https://doi.org/10.1111/jpim.12233>
11. Islam, J., Rahman, Z.: The impact of online brand community characteristics on customer engagement: an application of Stimulus-Organism-Response paradigm. *Telematics Inform.* **34**, 96–109 (2017). <https://doi.org/10.1016/j.tele.2017.01.004>
12. Brodie, R.J., Ilic, A., Juric, B., Hollebeek, L.: Consumer engagement in a virtual brand community: an exploratory analysis. *J. Bus. Res.* **66**, 105–114 (2013). <https://doi.org/10.1016/j.jbusres.2011.07.029>
13. Muñiz Albert, J., O'Guinn, T.: Brand community. *J. Consum. Res.* **27**, 412–432 (2001). <https://doi.org/10.1086/319618>
14. Merz, M.A., He, Y., Vargo, S.L.: The evolving brand logic: a service-dominant logic perspective. *J. Acad. Mark. Sci.* **37**, 328–344 (2009). <https://doi.org/10.1007/s11747-009-0143-3>
15. Muniz, A.M., Jr., O'Guinn, T.C.: Brand community. *J. Consum. Res.* **27**, 412–432 (2001). <https://doi.org/10.1086/319618>
16. Leigh, T.W., Peters, C., Shelton, J.: The consumer quest for authenticity: the multiplicity of meanings within the MG subculture of consumption. *J. Acad. Mark. Sci.* **34**, 481–493 (2006). <https://doi.org/10.1177/0092070306288403>
17. Steinhoff, L., Arli, D., Weaven, S., Kozlenkova, I.V.: Online relationship marketing. *J. Acad. Mark. Sci.* **47**, 369–393 (2019). <https://doi.org/10.1007/s11747-018-0621-6>
18. Ransbotham, S., Kane, G.C.: Membership turnover and collaboration success in online communities: explaining rises and falls from grace in Wikipedia. *MIS Q. Manag. Inf. Syst.* **35**, 613–627 (2011). <https://doi.org/10.2307/23042799>
19. Pongsakornrungsilp, S., Schroeder, J.E.: Understanding value co-creation in a co-consuming brand community. *Mark. Theory* **11**, 303–324 (2011). <https://doi.org/10.1177/1470593111408178>
20. Wieland, H., Koskela-Huotari, K., Vargo, S.L.: Extending actor participation in value creation: an institutional view. *J. Strateg. Mark.* **24**, 210–226 (2016). <https://doi.org/10.1080/0965254X.2015.1095225>

21. Boudreau, K.J., Lakhani, K.R.: Using the crowd as an innovation partner. *Harv. Bus. Rev.* **91**(60–9), 140 (2013)
22. Mahr, D., Lievens, A., Blazevic, V.: The value of customer cocreated knowledge during the innovation process. *J. Prod. Innov. Manag.* **31**, 599–615 (2014). <https://doi.org/10.1111/jpim.12116>
23. Kjellberg, H., Storbacka, K., Akaka, M., et al.: Market futures/future markets: research directions in the study of markets. *Mark. Theory* **12**, 219–223 (2012). <https://doi.org/10.1177/1470593112444382>
24. Ind, N., Iglesias, O., Schultz, M.: Building brands together: emergence and outcomes of co-creation. *Calif. Manag. Rev.* **55**, 5–26 (2013). <https://doi.org/10.1525/cmr.2013.55.3.5>
25. De, J.J.P.J., De, B.E.: Innovation lessons from 3-D printing. *MIT Sloan Manag. Rev.* **52**, 43–52 (2013)
26. Cova, B., Pace, S.: Brand community of convenience products: new forms of customer empowerment—the case “my Nutella The Community.” *Eur. J. Mark.* **40**, 1087–1105 (2006). <https://doi.org/10.1108/03090560610681023>
27. Schau, H., Arnould, E., Arnould, J., Muñiz Albert, J.: How brand community practices create value. *J. Mark. Am. Mark. Assoc. ISSN* **73**, 30–51 (2009). <https://doi.org/10.1509/jmkg.73.5.30>
28. Antorini, Y.M., Muñiz, A.M.: The benefits and challenges of collaborating with user communities. *Res. Technol. Manag.* **56**, 21–28 (2013). <https://doi.org/10.5437/08956308X5603931>
29. von Hippel, E.: Horizontal innovation networks—by and for users. *Ind. Corp. Change* **16**, 293–315 (2007). <https://doi.org/10.1093/icc/dtm005>
30. Sawhney, M., Verona, G., Prandelli, E.: Collaborating to create: the internet as a platform for customer engagement in product innovation. *J. Interact. Mark.* **19**, 4–17 (2005). <https://doi.org/10.1002/dir.20046>
31. Di Gangi, P.M., Wasko, M.: Steal my idea! Organizational adoption of user innovations from a user innovation community: a case study of Dell IdeaStorm. *Decis. Support Syst.* **48**, 303–312 (2009). <https://doi.org/10.1016/j.dss.2009.04.004>
32. Akman, H., Plewa, C., Conduit, J.: Co-creating value in online innovation communities. *Eur. J. Mark.* **53**, 1205–1233 (2019). <https://doi.org/10.1108/EJM-12-2016-0780>
33. Essamri, A., McKechnie, S., Winklhofer, H.: Co-creating corporate brand identity with online brand communities: a managerial perspective. *J. Bus. Res.* **96**, 366–375 (2019). <https://doi.org/10.1016/j.jbusres.2018.07.015>
34. Pantano, E., Priporas, C.V., Dennis, C.: A new approach to retailing for successful competition in the new smart scenario. *Int. J. Retail. Distrib. Manag.* **46**, 264–282 (2018). <https://doi.org/10.1108/IJRDM-04-2017-0080>
35. Corbin, J.M., Strauss, A.: Grounded theory research: procedures, canons, and evaluative criteria. *Qual. Sociol.* **13**, 3–21 (1990). <https://doi.org/10.1007/BF00988593>
36. Creswell, J.W., Hanson, W.E., Clark Plano, V.L., Morales, A.: Qualitative research designs: selection and implementation. *Couns. Psychol.* **35**, 236–264 (2007). <https://doi.org/10.1177/0011000006287390>
37. Taylor, S.J., Bogdan, R., DeVault, M.: Introduction to Qualitative Research Methods: A Guidebook and Resource—Steven J. Taylor, Robert Bogdan, Marjorie DeVault - Google Livros (1984). https://books.google.com.br/books?hl=pt-PT&lr=&id=pONoCgAAQBAJ&oi=fnd&pg=PR11&dq=Introduction+to+Qualitative+Research+Methods:+A+Guidebook+and+Resource&ots=Qitjgy4E1U&sig=izE_apiWp3Rken_CoYsG1hKsFhI#v=onepage&q=Introduction+to+Qualitative+Research+Methods. Accessed 25 Mar 2021
38. Mayring, P.: Qualitative Content Analysis: Theoretical Background and Procedures, 365–380 (2015). https://doi.org/10.1007/978-94-017-9181-6_13
39. Stake, R.E.: Qualitative case studies. In: *The Sage Handbook of Qualitative Research*, 3rd edn, pp. 443–466. Sage Publications Ltd, Thousand Oaks, CA (2005)
40. Coffman KG, Odlyzko AM (2002) Chapter 2 - Growth of the Internet

41. Bartl, M., Kannan, V.K., Stockinger, H.: A review and analysis of literature on netnography research. *Int. J. Technol. Mark.* **11**, 165 (2016). <https://doi.org/10.1504/ijtmkt.2016.075687>
42. Langer, R., Beckman, S.C.: Sensitive research topics: netnography revisited. *Qual. Mark. Res.* **8**, 189–203 (2005). <https://doi.org/10.1108/13522750510592454>
43. Kozinets, R.V.: Netnography. *Int. Encycl. Digit. Commun. Soc.* 1–8 (2015). <https://doi.org/10.1002/9781118767771.wbiedcs067>
44. Hsieh, H.F., Shannon, S.E.: Three approaches to qualitative content analysis. *Qual. Health Res.* **15**, 1277–1288 (2005). <https://doi.org/10.1177/1049732305276687>

Brands at the Nexus of Organizations and Markets: How Organizational Structures and Technology Shape Marketing Practice



Miguel Magalhães and Fernando Pinto Santos

Abstract Recent research suggests that marketing has been losing influence on some company's strategic decision-making and has been increasingly employed at a tactical level. Additionally, technological advancements in media and communications have been pointed out as being drivers of this "tactification" of marketing. Within this phenomena, we are particularly interested in understanding how brand management is being regarded in large organizations. Thus, we set as purpose for our research to explore how brand management is currently being practiced in large organizations, with a particular concern for understanding their strategic role in these companies. To this end, we have developed a qualitative study, where data were mainly collected through in-depth interviews to top managers. We have studied how three large organizations structure the marketing function and what roles brands assume within their strategic endeavours. Our findings unveil important insights on the role of brands in the practice of marketing at both strategic and tactical levels, laying ground for future research.

Keywords Brand management · Marketing practice · Strategy · Technology

1 Introduction

Globalization and technological innovation brought changes to consumer's usage patterns, which have affected marketing practice in terms of budget resource allocation, forcing organizations to rethink how they practice and structure marketing [1, 2]. Additionally, marketing has been losing influence on companies decision-making processes in recent years [3, 4]. Some researchers argue that organizations

M. Magalhães (✉) · F. P. Santos
IPAM Porto, Porto, Portugal
e-mail: miguel.magalhaes@universidadeuropeia.pt

F. P. Santos
e-mail: fernando.santos@ipam.pt

F. P. Santos
UNIDCOM/IADE, Lisboa, Portugal

have been downplaying the contribution of marketing to their strategic management and rather regarding it as a tactical practice concerned essentially with communication and sales [5, 6]. Although these views cannot be generalized, they nonetheless reveal interesting dynamics on marketing management that need to be explored and understood, given that there is a lack of studies that address how marketing is being practiced by commercial organizations [7, 8]. In an age where change, globalization and technological innovation are more prominent than ever, the role of marketing in organizations continues to evolve, and more studies are needed to understand it in contemporary practices [9–11].

Within marketing, we have a particular interest in understanding brand management and the role of brands in the practice of strategy. Although some studies have suggested that brands can assume a central strategic role that informs the wider practices of organizations, becoming a reference point for management in and between organizations and their markets [e.g. 12, 13], still there is much to further study. Thus, we set out as purpose for our research to explore how brand management is currently being practiced in large organizations, with a particular concern for understanding their strategic role in these companies. We have studied three large organizations to gain an in-depth understanding on the practices of these companies.

2 Related Literature

Throughout the years, organizations have changed in the way they practice and structure marketing [1, 14]. The integration of information technology into contemporary marketing practice has drastically changed the field [9, 15]. Globalization and technology, driven by the mobile phone and the Internet, spawned new paradigms on marketing in the areas of online advertising, e-commerce, and social media [11]. Media usage patterns have undergone significant changes over the years, with people spending more time on interactive media (social media) than on traditional media (radio and print). These changes in usage patterns have affected marketing practice in terms of budget resource allocation, forcing organizations to rethink their marketing [2].

Authors believe that the marketing discipline should adopt a more strategic orientation within research and organizations [16, 17]. The reason why strategy is so important to marketing is because a strategic orientation results in maximum value for the organization and having a marketing department with a strategic role enhances an organization's performance [18, 19].

The essential terms of the strategy are related to marketing, as the company's marketing policy is the basis for planning all aspects of the company's activities. In this context, marketing can be divided into two levels: strategic and tactical [20]. Strategic marketing involves a continuous analytical process focused on decisions like strategy type (e.g. cost vs. differentiation), value proposition/positioning, and

market selection/targeting [21]. As for the tactical level, it consists of an active short-term planning process, majorly related to sales and communications [20]. Moreover, tactics are mostly associated with the marketing mix [21].

However, the marketing department has lost influence in the company's decision-making [3, 4]. It is claimed that marketers have an excessive propensity to employ marketing at an essentially tactical level, disregarding its importance and value as a contribution to the strategic management of commercial organizations [5, 6]. This propensity towards tactics became exacerbated in the last decade, with the explosive growth of digital, and the current use of marketing as a tool in this domain is often directed simply to generate sales, which further leads to some practitioners disregarding the strategic role of marketing [3, 4, 9].

In the past years, marketing has taken a more tactical role in organizations [9, 11]. Communication became a key element of marketing activities [22] and was maximized by technology, with advancements in advertising and information [23]. These technological advancements, along with globalization, led to the digital marketing era [11, 24], taking scholars and practitioners to define and practice marketing mostly as communication and promotion [24].

3 Methodology

Exploratory in nature [25], our study follows a qualitative approach to address how marketing is being practiced in organizations. We have chosen three large organizations, operating at a global scale, as the empirical context where data were collected. These organizations were chosen based on purposive sampling [26], having been considered as suitable within the goals of this research. Importantly, good access to key informants and internal documents were also part of the reasons for the choice. The empirical material consists of three semi-structured interviews to the marketing decision-makers of the organizations and more than 100 pages of internal documents. The chosen organizations operate in different industries, as shown in Table 1, which provides information about them.

The empirical data were first analysed in an open manner [27]. The interviews were transcribed and repeatedly read. The analysis of the documents followed the same process. With the repeated readings, general descriptive notes to highlight relevant aspects of the analysed material were created, within the framing of the

Table 1 Information about the interviewees and their organizations

Interviewee	Role	Company	Industry
Bruno Oliveira	Senior Brand Manager	Sumol + Compal	Food production
Bruno Rio	Area Manager	Sonae MC	Retail
Catarina Castro	Head of Marketing	Panidor	Food and beverages

research purpose. Overall, with these notes, we aimed to identify recurrent themes that allowed us to understand the phenomenon of study. Interpretive content analysis [28] was employed in this process and over time the notes evolved to codes through which we organized and refined the interpretation of the empirical material, and that ultimately led to the findings of this study, which are going to be presented in the next section.

4 Findings

The studied organizations structure their marketing activities into different brands. Each brand is positioned to a specific market segment and has a manager. In turn, all managers' report to a director—in the studied organizations, an Area Manager, and a Head of Marketing. The following excerpts from these two marketing decision-makers describe the brand-based structure of their organizations:

The marketing team is divided into mega departments (...) and starts to branch out into different teams (...) in my team, there are two functions: brand managers and content managers. (Bruno Rio).

In the [corporate] group, we have 4 units (...) [the marketing decision] is all centralized in myself and then we have the collective of brand managers, that is, the brand guardians, which respond to me (...) so it's me and 4 brand managers. (Catarina Castro).

These organizations use brand management as a way to develop a specific planning and decision-making for each brand and consequently each market segment, instead of standardizing their strategy across all brands. Thus, brand managers act as “brand guardians”, as one of the respondents told us. They become the main responsible actors for understanding their brand, market, and customers. The choice of this structure is related with customer-centricity, as one respondent explained clearly:

I see marketing as focusing on serving the customer in an increasingly holistic, efficient way and serving the customer first (...) Marketing is the ability of a brand to serve a customer holistically and efficiently. (Bruno Rio).

Furthermore, this organizational structure is also important in terms of innovation. One of the decision-makers, that is Head of Marketing, and has brand managers under her management, explained how functions encompass innovation in a brand-based organizational structure:

There is an area here that I had never done before, which is the innovation area. Up until the date of joining this company, I only had to add associated value. Now, I must deliver the opportunity (...) that is, bring another way of looking at the product, and that's what happened (...) I am much more focused on innovation than on added value. (Catarina Castro).

As the Head of Marketing oversees the brands and supports product development to create innovation (either to improve product differentiation in existing markets or to explore new markets), the functions of brand managers are mostly related to strategic planning and development, as is illustrated by the following empirical excerpt:

[The brand managers] have the product development pipeline and then all the strategy design. Targeting and product value proposition (...) brand positioning (...) from the development of the communication strategy to the actual implementation and when I say the implementation it's 360: multimedia, trade marketing and digital in coordinating with the various teams. (Bruno Rio).

[As a brand manager] you have all the analytical part (...) let's call it the brand finances. Then, the digital part (...), then communication and content (...) and brand activation. (Bruno Oliveira).

In all organizations, the brand managers had influence in two other marketing functions apart from strategic planning: market intelligence and marketing tactics. The first one was a joint work between the marketing and analytics departments, but the brand managers and decision-makers also had direct access to intelligence via data tools, as it is explained by one interviewee:

The entire analytics team that performs customer segmentation and analysis to increasingly optimize the promotional component and understand what megatrends are happening to boost the sales team. (Bruno Rio).

I live with a lot of data, so Kantar and Nielsen [data tools] are my best friends. (Catarina Castro).

It is important to note that the studied organizations seem to have a disproportionate focus towards tactics, specifically communication. Although the tactical functions were still part of the brand management responsibilities, communication seemed to be attributed to others, still within the marketing department. Additionally, a few of the operational communications functions were also outsourced, as explained by one of the respondents:

We are much more focused on the logic of communication and promotion. (Bruno Rio).

We have 3 people doing all the management of social media and newsletters because we still live off newsletters (...) it's fundamental. (Catarina Castro).

I am responsible for [profit] margins, volume, and sales (...) Everything related to PR, we have an agency for that. (Bruno Oliveira).

Overall, the function of brand management was responsible for integrating market intelligence, strategic planning, and marketing tactics. Not only brand managers influence strategy, but also have ownership of strategic planning and of the guidelines used to develop marketing tactics.

5 Discussion

Figure 1 brings together the findings of our research and is useful to discuss the insights of our study.

5.1 Strategy Versus Tactics: Not “Either-Or” but “Both-And”

In the studied organizations, brand managers use market insights to create and further develop the marketing strategy for the market segment addressed by their brand. These brand managers also manage the marketing tactics related to the brands. They are, as one respondent put it, “the brand guardians”: they have ownership for the strategy and tactics of their own brands. Thus, in the empirical contexts addressed in this research, marketing is clearly being practiced at both strategic and tactical dimensions. Interestingly, the weight of the tactical dimension is quite relevant in terms of the practices and resources involved (namely human resources), but does not imply a downplay, in any way, of the contribution that a marketing approach has in strategic decision-making.

Technology has accelerated the tactification of marketing, due to the advancements in media and communications tools that unlock many possibilities for marketing practice [7]. With increasing opportunities to communicate and reach consumers, almost at times on a one-to-one logic, marketing practices do focus heavily on tactical activities. However, in the studied organizations, the tactical

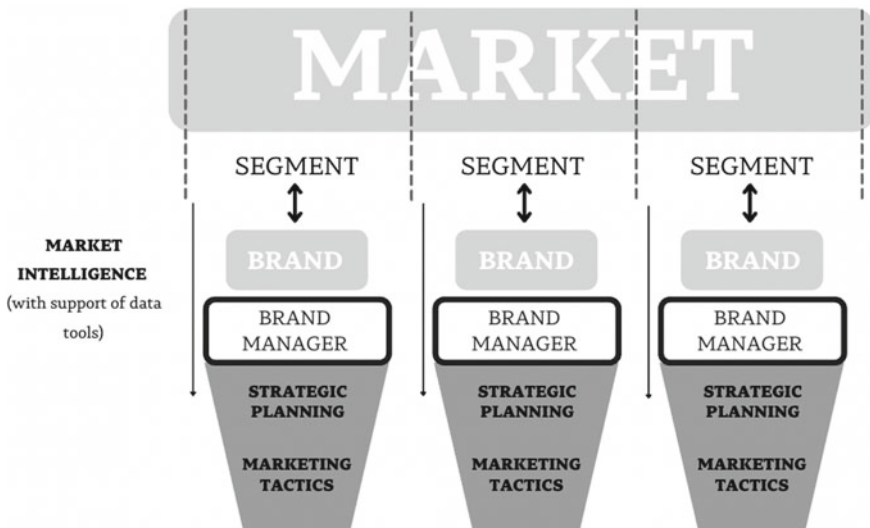


Fig. 1 Organizational structure based on brands

dimension is simply a part of the marketing activities and one that is quite embedded in strategic planning and decision-making. This is a finer-grained view on the discussions that put strategic and tactics as extremes in the practice of marketing [5].

The link between strategy and tactics can be tracked to the need of segmenting the market and creating brands for its segments: the brand-based structure helps simplify the market and act in a direct, focused manner in each segment. As brand managers become close to the reality of their market segment, their immediate function is to develop strategic planning based on market insights and translate the strategic guidelines into tactics. Therefore, the marketing activities are as strategic as they are tactical, considering that tactics and activities must follow the marketing strategy.

5.2 Brand as the Nexus Between Organizations and Markets

Our study further revealed that brands do assume a strategic role in the marketing department and the wider organization, becoming a point of reference for internal processes, at several levels. That is, brands seem to become lenses through which members in the organization regard the market. Interestingly, as it is through brands that the marketing activities are organized, they also become lenses in the flow of information from the market to the organization. Brands thus become the nexus of flow of information and activities, from and to markets and organizations. Technology plays a fundamental role since it is often the support of these flows and activities.

We thus have a close dynamic between brands as the marketing point of reference and technology as the apparatus that supports many of the activities related to brand management. For instance, market intelligence that is essential to feed the processes of strategic decision-making is collected with different digital technological solutions. On the other hand, most of the communication activities of the studied organizations do rely on digital technology. Thus, this tactical level is also embedded in technology and that has an impact on the communication activities themselves.

6 Conclusions

This study shows that brands assume a strategic role in organizations by being a reference point through which the marketing activities are organized. Importantly, organizations that have structured their marketing activities on brands, also develop their tactical activities through those brands, which is something that deserves more attention in future research. This finer-grained view on the dynamics between strategy and tactics in organizations is an avenue of research that is worth exploring, given the relevance of narrowing the gap between academic research and the work of practitioners.

Importantly, our research also reveals that brands become reference points at the nexus of organizations and markets, being lenses upon which the organizations regard the markets and on how the market is brought (through market intelligence) to the organization. This nexus is permeated by technology.

In overall terms, our study reveals that from the intertwining between brand management and technology, marketing practices emerge. These practices are shaped by the nature of the organizational structures and the different technological solutions that are employed in the studied organizations. These insights from our research raise important issues about the contemporary practices of marketing, which deserve more attention in future research.

References

1. Kirova, V.: Exploring the role of strategic marketing department. *J. Mark. Dev. Competitiveness* **11**(2), 27–38 (2017)
2. Kumar, V.: Evolution of marketing as a discipline: what has happened and what to look out for. *J. Mark.* **79**(1), 1–9 (2015)
3. Homburg, C., Vomberg, A., Enke, M., Grimm, P.: The loss of the marketing department's influence: is it really happening? And why worry? *J. Acad. Mark. Sci.* **43**(1), 1–13 (2015)
4. Verhoef, P., Leeflang, P., Reiner, J., Natter, M., Baker, W., Grinstein, A., Gustafsson, A., Morrison, P., Saunders, J.: A cross-national investigation into the marketing department's influence within the firm: toward initial empirical generalizations. *J. Int. Mark.* **19**(3), 59–86 (2011)
5. Contreras, F., Ramos, M.: What is marketing? A study on marketing managers' perception of the definition of marketing. *InForum Empresarial* **21**(1), 49–69 (2016)
6. Cornelissen, J.: Academic and practitioner theories of marketing. *Mark. Theory* **2**(1), 133–143 (2002)
7. Hackley, C.: *Marketing: A Critical Introduction*. Sage (2009)
8. Kumar, V.: Integrating theory and practice in marketing. *J. Mark.* **81**(2), 1–7 (2017)
9. Edwards, C., Bendickson, J., Baker, B., Solomon, S.: Entrepreneurship within the history of marketing. *J. Bus. Res.* **108**, 259–267 (2020)
10. Leeflang, P., Verhoef, P., Dahlström, P., Freundt, T.: Challenges and solutions for marketing in a digital era. *Eur. Manag. J.* **32**(1), 1–12 (2014)
11. Rangnekar, A.: The evolution of marketing and strategic thought. *Chetana's* **9**(1), 1–12 (2018)
12. Santos, F., Burghausen, M., Balmer, J.: Heritage branding orientation: the case of Ach. Brito and the dynamics between corporate and product heritage brands. *J. Brand Manage.* **23**(1), 67–88 (2016)
13. Santos, F., Benker, A., Lehtonen, M.: The role of design in the repositioning of a corporate brand. In: 11th European Academy of Design Conference 2016, Paris, France 22–24 (2015)
14. Webster, F.: The changing role of marketing in the corporation. *J. Mark.* **56**(4), 1–17 (1992)
15. Brady, M., Saren, M., Tzokas, N.: Integrating information technology into marketing practice—the IT reality of contemporary marketing practice. *J. Mark. Manag.* **18**(5–6), 555–577 (2002)
16. Hunt, S.: On rethinking marketing: our discipline, our practice, our methods. *Eur. J. Mark.* **28**(3), 13–25 (1994)
17. Webster, F.: Marketing IS management: the wisdom of Peter Drucker. *J. Acad. Mark. Sci.* **37**(1), 20–27 (2009)
18. Kumar, V., Shah, D.: Expanding the role of marketing: from customer equity to market capitalization. *J. Mark.* **73**(6), 119–136 (2009)

19. Wirtz, J., Tuzovic, S., Kuppelwieser, V.: The role of marketing in today's enterprises. *J. Serv. Manag.* **25**(2), 171–194 (2014)
20. Dubickis, D., Ginevičius, R.: The formation of structure of marketing department. *Mokslas–Lietuvos ateitis/Science–Future of Lithuania*, **6**(1), 17–24 (2014)
21. Morgan, N., Whitler, K., Feng, H., Chari, S.: Research in marketing strategy. *J. Acad. Mark. Sci.* **47**(1), 4–29 (2019)
22. Duncan, T., Moriarty, S.: A communication-based marketing model for managing relationships. *J. Mark.* **62**(2), 1–13 (1998)
23. Kitchen, P., Brignell, J., Li, T., Jones, G.: The emergence of IMC: a theoretical perspective. *J. Advert. Res.* **44**(1), 19–30 (2004)
24. Jain, E., Yadav, A.: Marketing and technology: role of technology in modern marketing. *IOSR J. Bus. Manag.* **19**(5), 49–53 (2017)
25. Patton, M.: *Qualitative Research and Methods: Integrating Theory and Practice*. Thousand Oaks, CA (2015)
26. Stake, R.: *The Art of Case Study Research*. Sage (1995)
27. Alvesson, M., and Sköldbberg, K.: *Reflexive Methodology: New Vistas for Qualitative Research*. Sage (2017)
28. Miles, M., Huberman, A.: *Qualitative Data Analysis: An Expanded Sourcebook*. Sage (1994)

The Impact of Brand Community on the Consumer Behavior: The Harley-Davidson Brand Community in Portugal Case



Joaquim Pratas, Maria João Vasconcelos, and José Luís Reis

Abstract The aim of this investigation is to understand the Harley-Davidson (HD) brand community members in Portugal and its consequences on their behaviors and brand loyalty. Throughout this work, it is mentioned how the HD brand developed the brand community, becoming an appealing brand for freedom lovers. This research focused on 429 members of HD brand community in Portugal. The methods used to analyze data were a descriptive data analysis, the chi-square independence tests, the chi-square automatic interaction detector (CHAID) model and the binary logistic regression. The results show that brand communities provide to its members the experience of integration with others, reinforcing brand loyalty. The merchandising and accessories purchase by brand community members is also reinforced.

Keywords Brand community · Harley-Davidson · Tribal marketing · Consumer behavior · Loyalty

1 Introduction

A brand community is a specialized, non-geographically bound community, based on a structured set of social relations among admirers of a brand [1]. It has high value for organizations' marketing, innovation management and customer relationship management [2], since brand community members influence each other, and their

J. Pratas (✉)

Universidade da Maia - ISMAI, Research Unit UNICES, Maia, Portugal

e-mail: jmpratas@ismai.pt

Research Unit GOVCOPP, Universidade de Aveiro, Aveiro, Portugal

M. J. Vasconcelos · J. L. Reis

University of Maia – ISMAI, Maia, Portugal

e-mail: a036329@ismai.pt

J. L. Reis

e-mail: jreis@ismai.pt

J. L. Reis

Research Unit LIACC—University of Porto, FEUP, Porto, Portuga

social interactions affect each individual relationship with the brand [3]. Generally, individuals participate in the communities in order to seek help to receive information which is specifically tailored to their needs [4], to establish emotional bonds [5] and increasingly organize brand communities based on their consumption, commitment and attachment to particular brands [6]. These brand communities members are brand lovers, as they are consumers who adopt the values and lifestyles that the brand represents. Brand communities exist based on the members who join with the same feeling of passion about a brand and based on what the brand has to offer them, establishing an affective relationship with the brand and not just a commercial relationship.

The consumer's involvement with a brand community has been decisive affecting brand sales [7], increasing brand loyalty and avoiding direct competition only based on price between brands [8].

This is the case of Harley-Davidson that will be studied, namely the brand community in Portugal and its member behaviors.

2 Literature Review

2.1 *Harley-Davidson (HD) Brand, Harley Owners Group (HOG) and 1%er Culture*

Harley-Davidson (HD) is the most evocative brand in the world and has remained iconic over the years [9], although no national television advertising and few press campaigns were made along its history [10]. Over the years, HD brand success has been related with the marketing techniques used, namely the definition of freedom, individualism and rebellion for its identity values. These marketing techniques led to brand loyalty that has achieved that some HD brand consumers tattooed the brand logo on their bodies as a form of expressing their passion, loyalty and commitment with the brand [11].

Reference [12] identified the characteristics of a HD's brand community, called Harley Owners Group (HOG). In this brand community, there is a subculture generated by the relationship between consumers and the brand and an ethos that is shared with the acculturation and hierarchy patterns, with some degree of marginalization and outlaw culture.

The HOG emerged in 1914 when HD owners decided to create an exclusive group for HD motorcycle buyers, so all new buyers of a HD motorcycle receive bi-monthly magazines, a portfolio from HOG, and are also invited to cafes in the concessionary stores and tours. HOG was then created with the aim of uniting people who loved HD brand, as members of this community adopt brand values, which are related with lifestyle and family belonging feelings.

HOG also sponsored races and events and developed branded merchandise such as apparel and accessories that could be bought by members. This strategy also allowed attracting new customers to HD products.

However, HOG had some members that damaged the brand image, namely the motorcycle clubs' members which are known as outlaw bikers. Generally, they appear on top of their bikes, using chains, sunglasses and oil-stained Levis, ready to go in action, with long hair blowing in the wind, beards and ribbons on the forehead flying, earrings, armpits exposed, chains and Harleys transformed with shining chrome [13]. These outlaw bikers are known as the 1%er culture, because after a rally event in 1943 occurred in Hollister, some bikers invaded the city, looted and destroyed stores. The American Motorcyclist Association (AMA) said that 99% of bikers were good men and within the law, while a minority of 1% was related with banditry and outlaws. This AMA's expression spread, and the outlaw bikers appropriated the term "1%er". Even today, many motorcycle clubs wear an embroidered "1%er" symbol to be identified as the members who belong to a restricted group of outlaws. These members defend the HD brand and their "brothers" in the brand community fiercely. Other groups related with marginality were created, namely the Bandits, the outlaws, the Pagan's, the Mongols, the Vagos and the Hells Angels. All these groups are known for their violent acts and criminal attacks but have an extremely rigid hierarchy within their organizations. The Hells Angels are the oldest and most famous international club in the world, where their motto is "*When we do it right, nobody remembers. When we make mistakes, no one forgets*". They are represented as a modern-day legend and a free spirit, becoming a camaraderie and freedom icon. What all motorcycle clubs have in common is that they ride custom motorcycles and HD brand, wear leather vests with the club's crest stamped on the back.

But with the HD's image linked to the outlaws pervading Americans' minds, the brand needed to find a way to deal with that image and assess its impact. They realized, then, that the new image of HD was not entirely negative, and that the outlaws' image created a big impact, having even increased company sales and "*exploited the rebel within its customers, mostly in males*" [10]. The company began to accept the two customer segments personalities and sell the brand through both directions. Today the HD brand portrays its brand not just as a motorcycle, but as a lifestyle icon, where the freedom value attracts both customer segments.

Today, the Harley Owners Group is made up of over one million members, with HOG being the largest motorcycle organization to be sponsored by the HD brand. The HOG community helped to invigorate the brand, so that it was no longer associated with marginality, banditry and rebellion. With HOG, the HD brand earned status, and the purchaser of a branded motorcycle is allowed to take part in an exclusive member group, where he can always feel at home, by participating in meetings and events. These HOG and "1%ers" tribes allowed the brand to acquire positive and negative components for its reputation. This duality has increased the brand's sales over the years, through the sale of motorcycles and merchandising. It should be noted that the customers' adhesion to the brand comes from the fact that the HD brand is

inserted in groups and communities, and consumers belonging to these groups end up relating to other consumers, making lasting friendships and generating a sense of belonging.

2.2 HOG Members' Consumer Behavior

Reference [14] stated that consumers that perceived the product effects relevancy tend to have a personal relationship with the products and high levels of involvement with the brand. Reference [15] also stated that consumers satisfy their freedom and self-definition needs through meaningful goods purchase and utilization, such as HD products.

Reference [16] identified two customer segments for HD product buyers: consumers with high monetary income (such as lawyers and doctors) who are able to buy motorcycles from this brand and consumers known as "outlaws". These consumers are mostly integrated in motorcycle clubs. Other identified characteristics are related with gender (about 90% are male consumers), race (mainly Caucasian), age (with a prevalence of members between 36 and 59 years), and with having tattoos, usually married with children, with casual or Rock N'Roll style clothing, and with musical taste linked to Rock N'Roll and Blues.

HD brand community members are also known for their festive get-togethers and make their presence felt at various events related to motorcycling and rock n' roll. Events and fellowships are sacred to the members of this cult, just like family fellowships. HD brand cult is the perfect combination between travel and the biker world. These gatherings give bikers the feeling of freedom. For the members of this cult's community, buying a motorcycle and riding a motorcycle are not enough for members to feel satisfied and fulfilled. For them, it is necessary to participate in events and gatherings, as only then will they meet the values of the HD brand, whose main values are friendship, companionship and freedom. For HD, the ideology is freedom and lifestyle [17], and many consumers identify with this same ideology being brand devotees. The brand also creates much of the "product" promoting the core values of personal freedom, machismo, patriotism and American heritage [18].

The motorcycle is just an entry to one of HD's many groups and brand communities. HD manages the customer relationship through a customer relationship marketing (CRM) program that involves components other than the main product itself, personalized services, brand merchandising, collectibles, etc.

In other words, the brand has such strong representation and notoriety that it transposes what is the company's core product. The motorcycle can be a ticket to the HD family. However, it takes more than that to stay in this community. The brand is crafted and communicated in such a way that consumers are able to mirror its values. This relationship between the company and the consumer is so strong that motorcycle owners tend to use the brand name and logo on their clothes, and in some cases, they even tattoo the brand symbol as a way to seal a commitment and a lifelong connection.

3 Proposed Conceptual Model, Research Questions and Methodology

3.1 Proposed Conceptual Model and Research Questions

Based on the existing literature, a conceptual model was proposed, and research questions were formulated in order to answer them in this research.

Individuals that belong to HD brand communities are the main consumers of the products, merchandising and accessories that brand sells. Thus, we want firstly to identify main characteristics of these individuals in terms of demographic and behavioral segmentation.

Purchasing a HD motorcycle is the “entry ticket” to the community with some status, but then investing in merchandising and its accessories seems important to assure brand loyalty and evolve in the brand community using symbolic icons.

Finally, the purchases of motorcycle, product and accessories can evolve to behaviors that are related with loyalty of brand members toward the brand. The proposed conceptual model is expressed in Fig. 1.

To the best of our knowledge, based on this conceptual model, the following research questions merit further investigation:

- Q1: What is the relationship between brand community members characteristics and the purchase of a HD motorcycle?
- Q2: What is the relationship between brand community members characteristics and the purchase of HD merchandising and accessories?
- Q3: What is the relationship between having a HD motorcycle and purchasing HD merchandising and accessories?
- Q4: What is the relationship between purchasing a HD motorcycle and loyalty behaviors toward HD brand?
- Q5: What is the relationship between the purchase of HD merchandising accessories and loyalty behaviors toward HD brand?

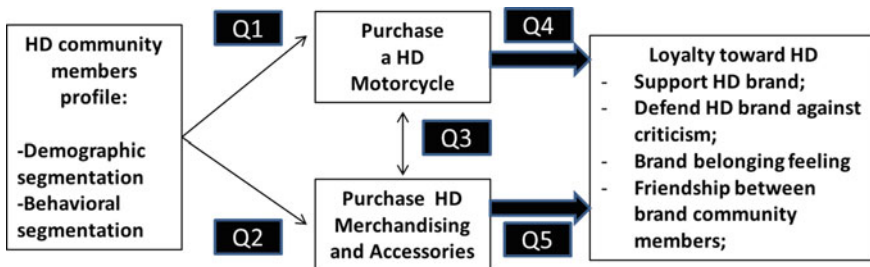


Fig. 1 Proposed conceptual model

3.2 Sample, Data Collection Method and Statistical Techniques Used

To answer these research questions, some procedures usually applied in quantitative researches were followed.

The selected sample for this research is based on 429 brand community members based on Portugal. Is a non-probabilistic and multistage sample, because first a sample of HD brand community members in Portugal who received the link to the online survey was chosen and then a voluntary sampling because brand community members took part and answer by themselves. The online survey that allowed the data collection had 26 questions. The questions used on the survey were based on work of [19] and preliminary exploratory interviews made with some HD brand community members in Portugal.

After collecting the data, a descriptive analysis of the results was carried using SPSS.

To assess the independence of variables, chi-square independence tests were carried out. In chi-square independence tests since the variables are nominal, cross-tabulation was performed to quantify the degree of relationship of the variables and the chi-square independence test. In the chi-square test, if the significance is less than 0.05 (with a 95% confidence interval), this does not mean that there is a cause and effect, but rather a relationship between the variables. If the significance is greater than 0.05, the variables are not related.

Then, other statistical methods, such as chi-square automatic interaction detector (CHAID) and the binary logistic regression methods, were used to analyze some of the proposed research hypothesis.

The chi-squared automatic interaction detection (CHAID) analysis generates predictive models as well as output diagrams by dividing a population into mutually exclusive subgroups after screening out several independent variables [20–22]. Thus, the dependent or explained variables are defined, and the independent or explanatory variables that most differentiate the results obtained are also identified.

A binary logistic regression was also developed. The objective of a logistic regression model is to fit a logit model relating the expected value of one binary dependent variable, to the actual values of several explanatory variables, in order stochastically to estimate the unknown parameters [22, 23].

4 Results

Firstly, a descriptive analysis of the data was performed. The relative frequencies for each survey question are presented in Table 1.

The sample of HD brand community members is mainly constituted by males, between 41 and 50 years old, married or similar, that dress in casual style. They prefer

Table 1 Relative frequencies of each question in the survey

Question	Relative Frequencies (%)
What is your gender?	Male (90.91%); female (9.09%)
What age group are you in?	Less than 30 (12.8%); between 31 and 40 (27.7%); between 41 and 50 (43.8%); between 51 and 60 (12.8%); More than 60 (2.8%)
What is your marital status?	Single (18.4%); married or similar (72.5%); divorced (7.9%); widower (0.9%)
What style of clothing do you wear?	Formal style (10.4%); Rock and Roll style (21.3%); casual style (57.2%); sporty style (11.1%)
What is your favorite music style?	Rock (23.4%); metal (13.8%); heavy metal (10.3%); pop (10.1%); other (37.4%)
What kind of space related with HD brand do you visit most frequently?	Motorcycle meetings and events (54.2%); Rock and Roll bars (32.4%); I do not visit any (13.3%)
Do you have tattoos?	Yes (55%); No (45%)
Do you have a HD motorcycle?	Yes (89.5%); No (10.5%)
What are your favorite HD motorcycle models?	Softail (16.4%); heritage (13.2%); fat boy (12.3%); iron (10.0%); other (48.1%)
In your opinion, which attributes do HD motorcycles have?	Freedom (32.2%); style (32.1%); power (12.6%); comfort (10.4%); other (less than 10%)
Do you feel that your motorcycle is an extension of you?	Yes (78.3%), maybe (13.1%), No (8.6%)
Is HD brand important for you?	Yes (87.9%); no (12.1%)
What do you feel, when someone criticizes HD brand?	I try to perceive the critics (48%); indifference (31.2%); Disagreement (14.2%); other options (less than 3%)
Does HD brand reflect you as a person?	Yes (54.6%); no (45.4%)
Do you feel that you would lose a part of your identity, if HD brand disappear?	Yes (35%); maybe (21%); no (44%)
Have you bought some HD merchandising?	Yes (80.2%); no (19.8%)
Do you think that HD users have more than one product of HD brand?	Yes (73.7%); maybe (18.7%); no (7.7%)
Do you think that HD users establish an emotional bond with brand?	Yes (88.8%); maybe (10.5%); no (0.7%)
Do you think HD owners and users are committed with the brand?	Yes (53.2%); maybe (26.8%); no (20.0%)
When a person uses something from the Harley-Davidson brand, do you identify with that person?	Yes (60.6%); no (39.4%)
Do you agree with the statement: "People who use a brand differ from those who don't"?	Yes (65.3%); no (34.7%)

(continued)

Table 1 (continued)

Question	Relative Frequencies (%)
Do you always disagree with someone prefers a competitive brand of HD?	Yes (89.7%); no (10.3%)
Does HD fit with you in ways that no other brand can?	Yes (69.5%); no (30.5%)
Do the friendships you have with other HD users mean a lot to you?	Yes (65.3%); no (34.7%)
Do you see yourself as a member of the HD brand community?	Yes (73.4%); no (26.6%)
Do you feel like an HD brand owner?	Yes (59.7%); no (40.3%)

several music styles such as rock, metal, heavy metal and pop music. The majority has tattoos and visits frequently motorcycle meetings and events.

89.5% of brand community members have a HD motorcycle, 80.2% have bought HD merchandising and accessories, and the most cited attributes of HD products are freedom and style.

The majority of inquired brand community members feel that HD motorcycle is an extension of them, important to them, reflects them as persons, fits with them as no other brand can do, and that when someone uses some product of HD they feel identified with him and that utilization differentiates that person.

It is also interesting to note that most respondents think that owners and users of HD brand establish an emotional bond and are committed with the HD brand.

They also feel that friendships with other HD users are important to them, and genuinely feel as HD brand owners and brand community members.

4.1 Relationship Between Brand Community Members Characteristics and the Purchase of a HD Motorcycle

To establish the relationship between brand community members characteristics and the purchase of a HD motorcycle, a chi-square analysis was performed. All the cross tabulations that have *P*-values less than $\alpha = 0.05$ are presented because it seems that there is a relationship between variables, so H_0 is rejected, as there is no independence between variables. These variables are expressed in Table 2. Only the favorite music Metal is negatively associated with purchasing a HD motorcycle. The likelihood of

Table 2 Brand community members characteristics associated with purchasing a HD motorcycle

Brand community members characteristics	Purchase a HD motorcycle
What is your favorite music style? Metal	0.966 (<i>P</i> -value = 0.034)

the combination involving the purchasing a HD motorcycle and Metal favorite music style is 0,966 times than the statistical non-association situation.

In all the others, *P*-values > 0.05, meaning that the distribution of the frequencies of two variable conditions was statistically non-associated.

4.2 Relationship Between Brand Community Members Characteristics and the Purchase of HD Merchandising and Accessories

All the cross tabulations that have *P*-values less than $\alpha = 0.05$ are presented because it seems that there is a relationship between variables. Thus, it seems that age groups, marital status, the frequent visits to rock and roll bars (with a likelihood of the combination involving the purchasing a HD merchandising and accessories and this variable is 1.06 times higher than the statistical non-association situation), motorcycle meetings and events (with a likelihood 1.05 times higher than non-association situation) and the perception of Freedom as an attribute related with HD brand are associated with purchasing HD merchandising and accessories (with a likelihood 1.04 times higher than non-association situation).

For age groups and marital status, the likelihood is not presented because we have several interval hypothesis, but it follows the same conclusions presented in the relative frequency analysis (Table 3).

Table 3 Brand community members characteristics associated with purchasing HD merchandising and accessories

Brand community members characteristics	Purchase HD merchandising and accessories
What age group are you in?	(-) (<i>P</i> -value = 0.001)
What is your marital status?	(-) (<i>P</i> -value = 0.002)
What kind of space related with HD brand do you visit most frequently? Rock and Roll Bars	1.06 (<i>P</i> -value = 0.028)
What kind of space related with HD brand do you visit most frequently? Motorcycle meetings and events	1.05 (<i>P</i> -value = 0.001)
In your opinion, which attributes do HD motorcycles have? Freedom	1.04 (<i>P</i> -value = 0.000)

Table 4 Having HD motorcycle and purchasing HD merchandising and accessories

To have a HD motorcycle	Purchase HD merchandising and accessories
Do you have a HD motorcycle?	1.05 (<i>P</i> -value = 0.000)

Table 5 Having a HD motorcycle association with loyalty behaviors toward HD brand

Loyalty behaviors toward HD brand	To have a HD motorcycle
Does HD fit with you in ways that no other brand can?	1.04; (<i>P</i> -value = 0.000)
Do you see yourself as a member of the HD brand community?	1.03; (<i>P</i> -value = 0.004)
Do you feel like a HD brand owner?	1.07; (<i>P</i> -value = 0.000)
Do you feel that you would lose a part of your identity, if HD brand disappear?	(-); (<i>P</i> -value = 0.050)

4.3 Relationship Between Having a HD Motorcycle and Purchasing HD Merchandising and Accessories

The likelihood of the combination involving having HD motorcycle and purchase of merchandising and accessories is 1.05 times than the statistical non-association situation, with a *P*-value equal to 0.000 (less than the 0.05 cut off value) (Table 4).

4.4 Relationship Between Having a HD Motorcycle and Loyalty Behaviors Toward HD Brand

All the cross tabulations that have *P*-values less than $\alpha = 0.05$ are presented because it seems that there is a relationship between variables. Thus, these variables are associated with having a HD motorcycle. In all the others, *P*-values > 0.05, meaning that the distribution of the frequencies of two variable conditions was statistically non-associated (Table 5).

4.5 Relationship Between the Purchase of HD Merchandising and Accessories and Loyalty Behaviors Toward HD Brand

To analyze the relationship between the purchase of merchandising and accessories and loyalty behaviors toward HD brand, a CHAID analysis was also performed. The results are represented by a CHAID dendogram. The variables that differentiate the purchase of HD merchandising and accessories are the answers to the survey questions “Do the friendships you have with other HD users mean a lot to you?” and “When a person wears something HD brand, do they identify with that person?”.

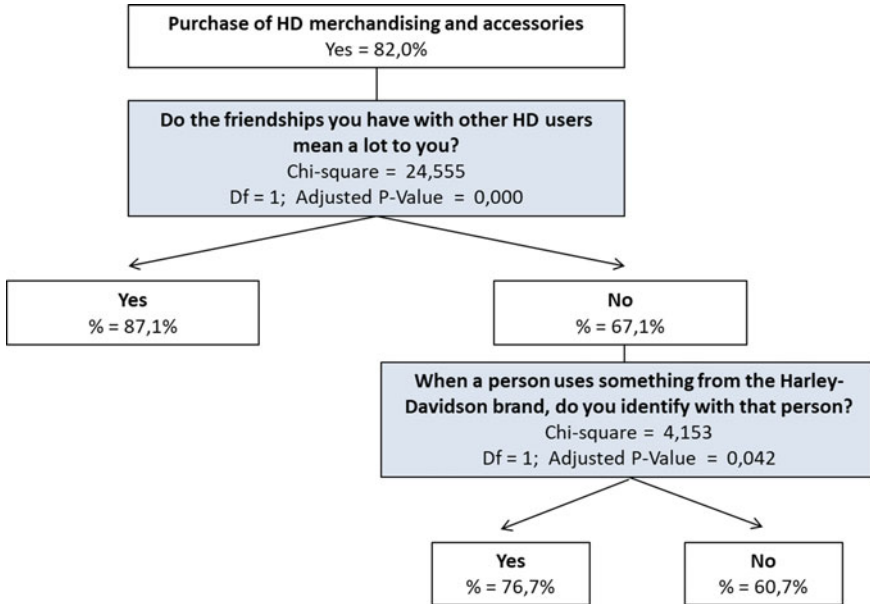


Fig. 2 CHAID dendrogram

In the entire sample, we can verify that 19.8% (85 members) did not buy merchandising and accessories, and 80.2% (with 344 members) bought.

Of the respondents who answered “Yes” to this question, 87.1% answered “Yes” to “Do the friendships you have with other HD users mean a lot to you?”.

On the other hand, of the respondents who answered “No” to this question, 67.1% answered “Yes”.

Therefore, it seems to be a positive relationship between the importance of friendships among users of the HD brand and identification with other people through the use of the HD brand and the purchase of HD merchandising and accessories (Fig. 2).

A binary logistic regression was also performed, in order to explain the purchase of HD merchandise and accessories. The results are showed in Table 6.

This regression has achieved 81.6% of correlation between predicted and observed values. The variable “Do the friendships you have with other HD users mean a lot to you?” has a *P*-value < 0.05. As Exp (B) is higher than 1, if friendships are important, the likelihood of purchasing HD merchandise and accessories increases.

5 Discussion and Future Research Directions

This investigation contributes to the study of brand community members’ behaviors, because nowadays the market is very competitive and volatile, and consumers

Table 6 Logistic regression model coefficients

	Purchase HD merchandise and accessories		
	Exp(B)	Wald	P-value
Do you feel that your motorcycle is an extension of you?	1.121	.322	.570
Is HD brand important for you?	1.691	1.941	.164
Does HD brand reflect you as a person?	.713	.914	.339
Do you feel that you would lose a part of your identity, if HD brand disappears?	1.211	.876	.349
Do you think that HD users establish an emotional bond with brand?	1.471	1.245	.265
Do you think HD owners and users are committed with the brand?	.850	.798	.372
When a person uses something from the Harley-Davidson brand, do you identify with that person?	1.216	.387	.534
Do you agree with the statement: "People who use a brand differ from those who don't"?	1.725	2.636	.104
Do you always disagree with someone prefers a competitive brand of HD?	1.714	.888	.346
Does HD fit with you in ways that no other brand can?	1.571	1.922	.166
Do the friendships you have with other HD users mean a lot to you?	1.938	5.007	.025
Do you see yourself as a member of the HD brand community?	1.036	.011	.916
Do you feel like an HD brand owner?	1.209	.479	.489
<i>-2LL</i>	380.940		
<i>Cox&Snell R²</i>	.102		
<i>Nagelkerke R²</i>	.162		
<i>% correlation between predicted and observed values</i>	81.6%		

are increasingly non-loyal, which makes the existence of brand communities very important to businesses.

Thus, brand communities arouse interest for those working in marketing and management areas. Harley-Davidson is a case study of successful brand community, so other brands could learn how to manage a brand community and understand what makes important to its members to live according to brand values and identity and increase sales of main products (such as the motorcycles for HD brand) and other merchandising and accessories.

In this research, that studied the HD brand community in Portugal, the behaviors of brand members were identified. The relationship brand community and the purchase of merchandising and accessories of the brand seems to be the most important variable in order to access the members' loyalty through periods, because having a motorcycle is "only" the entry door to the brand community. The best customers, who buy more products, tend to be the most engaged in brand communities. They are also the members that value more the friendship bonds with other brand community members.

Some future research directions can also be referred. This research can be done in other countries, and other Harley-Davidson brand communities and online brand communities can also be analyzed. Certainly, the interaction between brand community members in online environments and social media can add new perspectives to brand community behaviors.

References

1. Muniz, A.M., Jr., O'Guinn, T.C.: Brand community. *J. Consum. Res.* **27**(4), 412–432 (2001)
2. Zaglia, M.E.: Brand communities embedded in social networks. *J. Bus. Res.* **66**, 216–223 (2013)
3. McAlexander, J.H., Schouten, J.W., Koenig, H.F.: Building brand community. *J. Mark.* **66**(1), 38–54 (2002)
4. Dholakia, U.M., Bagozzi, R.P., Klein, P., Lisa, R.: A social influence model of consumer participation in network- and small-group-based virtual communities. *Int. J. Res. Mark.* **21**(3), 241–263 (2004)
5. Zhou, Z., Zhang, Q., Su, C., Zhou, N.: How do brand communities generate brand relationships? Intermediate mechanisms. *J. Bus. Res.* **65**(7), 890–895 (2012). <https://doi.org/10.1016/j.jbusres.2011.06.034>
6. Solomon, M., Bamossy, G., Askegaard, S., Hogg, M.K.: *Consumer Behaviour. A European Perspective*, 4th edn. Rotolito Lombarda, Italy (2010)
7. Noble, C., Adjei, T., Noble, S.: The influence of C2C communications in online brand communities on customer purchase behavior. *J. Acad. Mark. Sci.* **38**(5), 634–653 (2010)
8. Luedicke, K., Giesler, M.: *Brand Communities and Their Social Antagonists: Insights from the Hummer Case*. *Consumer Tribes*, pp. 275–295. Elsevier, Oxford (2007)
9. Frachetta, A.: Harley-Davidson: além de uma marca, um estilo de vida, from <https://www.estudioroxo.com.br/blogpulsar/harley-davidson-mais-que-uma-marca-um-estilo-de-vida/>. (2021)
10. Hill, S., Rifkin, G.: *Radical Marketing: From Harvard to Harley Lessons Learned from the Top Ten that Broke the Rules and Made It Big*. Harper Business, New York (1999)
11. Santos, E., Silva, D., Junior, S., Nascimento, C.: Comportamento do consumidor da comunidade de marca Harley-Davidson e a influência do self-expandido. *Remark: Revista Brasileira de Marketing*, **16**(1), (2017)
12. Shouten, J., Mcalexender, J.: Subcultures of consumption: an ethnography of the New Bikers. *J. Consum. Res.* 43–61 (1995)
13. Thompson, H.: *Hell's Angels*. Texto Editores (2007)
14. Peter, J.P., Olson, J.: *Comportamento do consumidor e estratégia de marketing*. Porto Alegre (2010)
15. McCracken, G.: *Cultura e consumo: uma explicação teórica da estrutura e do movimento do significado cultural dos bens de consumo*. São Paulo (2007)
16. França, C., Carvalho, E.: Profile of Motorcycles Harley-Davidson Consumers. *Revista Eletrônica de Ciências Empresariais* (2016)
17. Azevedo, M., Araújo, C.: A Marca Harley-Davidson no Brasil: o “Sonho” e os Entraves à Reprodução de uma Estratégia de Sucesso. *Encontro Nacional de estudos de consumo* (2010)
18. Dowling, G.: Customer relationship management: in B2C markets, often less is more. *Calif. Manage. Rev.* 87–103 (2002)
19. Taute, H., Sierra, J.: Brand tribalism: an anthropological perspective. *J. Prod. Brand Manage.* **23** (2014)
20. Kass, G.: An exploratory technique for investigating large quantitative of categorical data. *J. Appl. Stat.* **29**(2), 119–127 (1980)

21. Pratas, J.: Conteúdos de comunicação e de posicionamento das marcas destino turístico em brochuras e folhetos turísticos, PhD Thesis, Faculdade de Economia do Porto, Universidade do Porto (2014)
22. Brito, P., Pratas, J.M.: Tourism brochures: linking message strategies, tactics and brand destination attributes. *Tour. Manage.* **48**, 123–138 (2015). <https://doi.org/10.1016/j.tourman.2014.10.013>
23. Brito, P., Pratas, J.M.: Online brochures: relationship between message strategies and communicated attributes. *Anatolia* **27**(1), 47–57 (2016). <https://doi.org/10.1080/13032917.2015.1083208>

Luxury Fashion Brands' Website Strategies: A Study with Portuguese Designers



Sandrina Teixeira , Belém Barbosa , Cátia Ferreira, and José Luís Reis 

Abstract This article aims to identify the success factors for Website strategies of independent luxury fashion designers to communicate their brands and to foster lasting relationships with customers. The article includes an empirical study with 14 Portuguese luxury fashion brands designers that combine Website analysis and interviews with its managers. This study evidences the difficulties that Portuguese luxury fashion designers face in communicating the differentiating elements of a luxury fashion brand in digital channels, particularly product characteristics, luxury atmosphere, service personalisation, and the intimacy enabled by brick-and-mortar stores. The findings highlight the difficulties in strengthening online customer relationships, and those omnichannel strategies are essential to preserving luxury products' differentiating features and elements.

Keywords Fashion brands · Designers · Luxury · Website strategies · Online atmosphere

1 Introduction

Luxury is undeniably characterised by exclusivity, rarity, and superior quality. Historically, luxury was associated with wealth, power, and superfluous and non-essential

S. Teixeira

ISCAP, CEOS.PP, Porto Accounting and Business School, Polytechnic Institute of Porto, Matosinhos, Portugal

e-mail: sandrina@iscap.ipp.pt

B. Barbosa · C. Ferreira

School of Economics and Management, University of Porto, Porto, Portugal

J. L. Reis (✉)

University of Maia - ISMAI, Maia, Portugal

e-mail: jreis@ismai.pt

FEUP, LIACC - Artificial Intelligence and Computer Science Laboratory, University of Porto, Rua Dr. Roberto Frias, Porto, Portugal

desires [10]. However, modern luxury has a positioning beyond depravity or exclusivity, positioning itself to capture dreams and as a means of self-expression [3]. Luxury fashion brands have been quite reluctant to go digital and only recently have committed to digital strategies (De Acetis 2020), Hoang [7], with evident positive impacts on sales growth and brand equity and interesting opportunities to provide customer experience [1]. Despite the interesting contributions in the literature, the studies on this field often focus on well-established global brands. Consequently, research on the experiences and strategies of smaller Luxury Fashion Designer Brands (LFDB) from less prominent countries in the fashion business is still lacking.

Considering the gap found in the literature, the main aim of this article is to analyse the facilitators constraints faced by fashion designers in implementing e-commerce strategies and to point out their success factors. To reveal experiences and perspectives of LFDB, this article includes results from a study conducted in Portugal with 14 luxury brands, comprising content analysis of their Websites and interviews with their managers.

This article makes several contributions. It provides empirical data regarding a type of fashion luxury brand that the literature has so far disregarded. By comparing its findings with results from extant literature, particularly Marques [11], this article provides interesting insights on comparing Portuguese LFDB and international luxury fashion companies. This study demonstrates that the Website strategies conducted by small LFDB are distinct from the international brands, and for that reason, more research should address their challenges and strategies. Furthermore, this article provides relevant insights for academics and practitioners associated with digital strategies conducted by luxury fashion brands.

The remainder of the article is organised as follows. The next section offers a theoretical contextualisation of the topic provided by the main contributions in the literature regarding luxury brands' online strategies. Then, the article characterises the methodology adopted in the empirical study, followed by its main findings and discussion of results. The article's final section presents the main conclusions, highlighting their contributions and implications for managers. Suggestions for future research are also provided.

2 Literature Review

2.1 The Concept of Luxury

According to Roux and Flock [13], a luxury brand satisfies its consumers' symbolic needs and is distinguished from other brands by its imaginary, symbolic, and social value. Based on an exhaustive systematisation of the literature, Ko et al. [8] have defined five essential elements for any luxury brand:

- To be of high quality.

- Offering authentic value through the desired benefits, both functionally and emotionally.
- To have a prestigious image within the market, built on qualities such as craftsmanship and quality of service.
- Be worthy of possessing a premium price.
- Be able to inspire a deep connection or resonance with its consumers.

Ko et al. [8] pointed out that consumer evaluations and perceptions are crucial for a brand to be considered a luxury beyond premium prices and superior quality. Considering the irrefutable importance of digital channels as a source of information and their influence shaping consumers' perceptions and attitudes, it becomes increasingly essential that luxury fashion brands are online for enhanced opportunities to reach and interact with their customers.

2.2 The Atmosphere of Online Luxury

Rovai [14] explains in detail the demand that the entry into digital channels places on luxury brands, namely to demonstrate the authenticity of luxury products and recreate the comfort and overall experience provided at physical stores. Determinant aspects of the online atmosphere include Web design, the use of colour [16] and immersive environments [12]. The luxury atmosphere positively impacts the shopping experience, namely the emotion [16] and satisfaction [12, 14] felt by the user. However, [14] study demonstrates that omnichannel strategies luxury fashion brands are adopting include, for example, the creation of interactive online shops, conducting online events to promote brand engagement, and the creation of online apps. Beyond shopping touchpoints, online platforms are spaces where consumers live their online shopping experience, where they share know-how, trends, and lifestyle, seeking a connected shopping experience [14].

2.3 Online Strategies of Luxury Brands

Yan and Qu [17] recommend that luxury brands adopt pull strategies. They should attract consumers to seek the brand and not take it to consumers, as happens in the mass market. They also recommend choosing channels that preserve the exclusivity associated with luxury brands and foster relationships with their customers. Digital channels affect the development of luxury brands' strategies on several levels. Digital channels are an asset for luxury fashion brands, as online interactions between consumers, brands, bloggers, and sellers allow the brand to create social status associations through the easy sharing of information about what is trendy and stylish and through the dissemination of word-of-mouth communication about new trends [4]. However, they can also reduce the perception of scarcity by increasing proximity

to a broad group of potential consumers [4], leading many luxury fashion brands to recreate that distance in digital channels.

3 Method

Literature highlights the relevance of digital channels for luxury brands, pointing out aspects to be considered when adopting omnichannel strategies by luxury brands. However, the literature tends to privilege large global brands, thus neglecting many independent designer brands spread worldwide.

The main objective of this study is to identify the success factors for Website strategies of independent luxury fashion designers to communicate their brands and to foster lasting relationships with customers. This study explores and analyses fashion designers facilitators and constraints in implementing digital strategies. The research problem defined for this study is: What are the key elements of the omnichannel strategy conducted by LFDB?

This study adopted a content analysis approach to the research problem, comprising the analysis of brands' online platforms (Website and social media) and semi-structured interviews with LFDB managers.

3.1 Website Analysis

This study combined two analysis frameworks previously developed by extant literature. Thus, the analysis of the Websites was carried out using two analysis grids: Hansen's grid [5] and Teixeira and Machado's [15] digital visual merchandising grid—MVD.

The Websites of the brands/designers were consulted only once to collect data for analysis, whilst the social networks of the brands/designers were consulted during June 2019 to allow consideration of content that is shared in real-time (live streams) that has a duration of only 24 h (storeys) and paid advertisements also have cut-off dates.

Hansen's analysis grid was constructed to develop a theoretically consistent framework that identifies all relevant evaluation categories and subcategories for B2C Websites in the luxury fashion industry [6]. This grid posits 8 categories and 29 analysis variables.

The digital visual merchandising grid—MVD is intended for the analysis of the way products are presented in the virtual shop through 6 categories of variables and 34 variables of analysis [15].

To better understand the Portuguese LFDBs strategies, Hansen's grid and the digital visual merchandising grid—MVD were compared with Marques [11] results concerning international luxury fashion brands.

3.2 Interviews with LFDB Managers

To further explore the topic and validate the findings related to Website analysis, all LFDB analysed in this study was contacted and invited to participate in the interviews. Three out of the 14 brands considered in this study promptly agreed to participate (Josefinas, Carla Pontes, e Katty Xiomara). Interviews were conducted either in person or through the Internet during May 2019.

The semi-structured interview script was previously developed based on the contributions in the literature and was composed of three main parts. The first part comprised a set of questions about the brand strategy (e.g. What have been the significant changes at the strategic level of the brand). Next, the script included questions about the digital strategy, including the Website and social networks (e.g. How does the brand attract customers' attention or potential customers online? How does the brand keep in touch with its customers online?). In the third part of the script and conclusion, questions about digital challenges were included (e.g. What have been the main challenges encountered in the digital environment to date? What are the key points of the brand's online strategy?).

Content analysis is a set of techniques for analysing a brand's communications through the systematic and objective analysis of the content of its messages [2] to critically understand the meaning of the content of a brand's communications. In this study, linguistic analysis and mass communication analysis will be used.

The analysis of the interviews followed a categorical analysis in WebQDA software. According to [2], the categorical analysis allows the analysis of direct and straightforward speeches quickly and effectively, through the breakdown of the text into categories according to analogical regroupings.

To define the coding of the analysis, a first analysis of the interviews was carried out. The following categories were identified: online presence, current digital strategy, online challenges, strategic changes, omnichannel, product availability, luxury atmosphere, and characteristics of online luxury products/services, customisation, co-creation, secondary brand/more accessible product line, currency, promotional actions, value creation, consumer journey: attention, consumer journey: attraction, consumer journey: advice, consumer journey: action, consumer journey: advocacy, customer conversation/interaction, and online vs physical shop.

3.3 Sampling

The target population of the study was the Portuguese LFDB. Given the inexistence of statistical studies and rankings with all Portuguese luxury fashion brands and designers, and through online research, it was possible to collect Portuguese brands and designers on the Luxury Design Craftsmanship Website (visited on April 10, 2019), on the Portugal Fashion Website (visited on April 10, 2019). Next, an analysis of the brands and designers was carried out. Only those that work in the luxury

Table 1 Sample of Portuguese design luxury fashion brands

	Brand	Web site	Online store	Physical store	Facebook	Instagram	Twitter	Youtube	Pinterest	Other
1	Miguel Vieira	✓	✗	✓	✓	✓	✓	✓	✓	✗
2	Luís Buchinho	✓	✓	✓	✓	✓	✗	✗	✓	✗
3	Katty Xiomara	✓	✓	✓	✓	✓	✓	✓	✓	✗
4	Micaela Oliveira	✓	✗	✓	✓	✓	✓	✓	✓	✓
5	Luís Onofre	✓	✓	✓	✓	✓	✗	✓	✓	✗
6	Anabela Baldaque	✓	✗	✓	✓	✓	✗	✗	✗	✗
7	Diogo Miranda	✓	✓	✓	✓	✓	✗	✓	✗	✗
8	Egídio Alves	✓	✓	✗	✓	✓	✓	✓	✓	✗
9	Guava	✓	✓	✓	✓	✓	✗	✓	✗	✗
10	Carla Pontes	✓	✓	✓	✓	✓	✗	✓	✗	✗
11	Âme Moi	✓	✓	✗	✓	✓	✓	✓	✓	✗
12	Lemon Jelly	✓	✓	✗	✓	✓	✓	✓	✗	✗
13	Josefinas	✓	✓	✗	✓	✓	✓	✓	✓	✗
14	Nuno Baltazar	✓	✗	✓	✓	✓	✗	✗	✗	✓

fashion sector and are present online were selected, having been identified 17 brands. After analysing their respective product portfolios, it was found that three of these brands did not offer fashion products aimed at women. To make the sample more homogeneous, it was decided to remove the exclusively male brands. Consequently, the study sample consists of 14 Portuguese LFDB (Table 1).

All LFDB that comprises the sample has a Website, however, four still do not offer commercial transactions, and another four do not have a physical shop. It was also concluded that all are present on social networks, Facebook, and Instagram, followed by YouTube, Pinterest, and Twitter.

4 Results

4.1 Website Analysis—Hansen Framework

As shown in Table 2, the variables most used by Portuguese LFDBs in their Websites are content, connection, and communication. The least worked are collaboration, community, and customisation.

Table 2 Results from Portuguese LFDB Websites according to Hansen framework

Hansen's analysis grid		Total	%	Values reported by Marques [11] for international luxury fashion brands (%)
Content	Product information	10	71	100
	Brand history	12	86	87
	Corporate information	8	57	100
	Special promotions/campaigns	6	43	87
Community	Community on the site	0	0	20
	The community outside the site	5	36	93
Communication	Email registration	9	64	100
	About us/contacts	14	100	100
	Interview videos	0	0	40
Collaboration	Product design collaboration	0	0	0
	Allows feedback and comments	1	7	7
Connection	Links to other sites	13	93	60
	Microsites	0	0	33
	Syndication	8	57	100
Commerce	Online shop	0	0	0
	Commercial transaction	10	71	93
	<i>Links to other e-commerce sites</i>	2	14	0
	E-commerce assistant	7	50	87
	<i>Shoppable videos</i>	0	0	0
	<i>Social commerce</i>	10	71	0
Context	<i>Flash</i>	0	0	0
	Videos	7	50	93
	Animations	12	86	53
	360° visualisation	0	0	7
	Augmented reality	0	0	0
	3D Technology	0	0	0
Customisation	Product customisation	0	0	27
	Product personalisation	1	7	27
	User recognition	10	71	93

Marques [11] showed that international luxury fashion brands significantly use the several content variables considered in the Hansen framework (between 87 and 100%), which does not occur in the Portuguese companies analysed here. Notably, 43% of LFDB do not provide corporate information on their Websites, and only 43% include promotions.

It is in the macro community variable that Portuguese LFDB shows lower percentage values. Although 93% of the international luxury fashion brands have a community outside the Website [11], in the case of Portuguese LFDB, this happens only for 36% of the cases. It is essential to highlight that the communities outside the Portuguese LFDB Websites are in the social network Facebook. However, although some brands have a community in the social network, this one does not present activity.

Regarding the communication variables, all the Portuguese LFDB Websites and about us/contacts happened with the international luxury fashion brands studied by Marques [11]. However, none includes the Websites videos with interviews, which tends to be adopted by 40% of the international luxury fashion brands [11]. The macro collaboration variable is the one that presents the lowest values for Portuguese LFDB, confirming the trend amongst international brands identified by Marques [11].

In the macro variable connection, the Portuguese LFDBs present a higher percentage value than the International Luxury Fashion Brands in the variable linking to other websites: 93%, compared to the 60% reported by Marques [11]. The links to other Websites of the Portuguese Luxury Fashion Brands/Designers are mainly to the brand's social networks and the Web designer's Website. It should also be noted that no Portuguese LFDB presents microsite, something that is common to 40% of the international brands analysed by Marques [11].

Regarding the trade dimension, not all Portuguese LFDBs provide commercial transactions on their Website, and only 50% have an e-commerce assistant. In comparison, 87% of the international Luxury Fashion Brands already do so.

It should also be noted that at the date of the respective analyses, no Portuguese LFDB and none of the International Luxury Fashion Brands studied by Marques [11] have shoppable videos (videos that allow the purchase of products) and virtual shops. The same happened with technology to recreate the offline shopping experience in the digital environment, namely 3D technology and augmented reality. Although 7% of the International Luxury Fashion Brands provide 360° visualisation on their Website, Portuguese LFDB does so.

Moreover, Portuguese LFDB should invest more in videos since they present a percentage value much lower than International Luxury Fashion Brands. However, Portuguese LFDB shows more animations on their Websites than International Luxury Fashion Brands, 86% and 53%. Finally, no Portuguese Luxury Fashion Brand/Designer and International Luxury Fashion Brand uses Flash on their Website and all use HTML5.

According to Kotler et al. [9], the new product development strategy is co-creation in the digital economy. From the moment brands allow customers to personalise and customise their products online, they receive information regarding their tastes and preferences. On the other hand, consumers of luxury products expect the online shopping experience to equal the offline shopping experience [14]. If brands allow customisation and personalisation of their offline products, they should also make this service available online. However, no Portuguese LFDBs enable the customisa-

tion of their products on the Website and only one provides product customisation. Comparatively, four International Luxury Fashion Brands make these technologies available on their Websites.

Although 71% of the Portuguese Luxury Fashion Brands/Designers already do user recognition, it is essential to highlight that one of the brands that allow commercial transactions on its Website does not recognise the user. In comparison, 93% of the International Luxury Fashion Brands already identify the user on their Website.

4.2 Analysis of the Results of the Digital Visual Merchandising Grid—MVD

Considering the digital visual merchandising grid—MVD of Teixeira and Machado [15], it is verified that only one Portuguese brand does not have a front view of their products. In turn, all International Luxury Fashion Brands provide this variable on their websites.

Regarding the back view, two Portuguese brands do not have this variable on their Websites. In contrast, 93% of the International Luxury Fashion Brands allow the rear view of their products. The side view is less common in both, however, the International Luxury Fashion Brands (80%) present a percentage value higher than the Portuguese LFDB (71%).

Portuguese LFDBs present higher percentage values than International Luxury Fashion Brands [11] in the variables partial zoom, 2D longer view on the same page, 2D automatic change front and back, and 2D front and back on the same page. However, the Portuguese LFDBs do not allow multiple zooms, 2D change F/T cursor passage, whilst the International Luxury Fashion Brands already have these functionalities. No Portuguese LFDB and International Luxury Fashion Brand allow 3D more extensive views on separate pages, 3D automatic rotation, and 3D click-on rotation, features that would decrease the digital barrier between customer and product.

Beyond using technology, for product perception, it is equally crucial that luxury fashion brands provide the same shopping experience online and offline. For example, it is important to digitally transfer the assistant's service in the physical shop through the ease of selecting the desired colour, visualising the piece's wearability, and suggesting complementary pieces to the chosen one. Regarding the presentation of product colours, less than half of the Portuguese LFDBs present all the product colours on the same page, the remaining ones choose to present products in different colours on separate pages. Comparatively, only 20% of the International Luxury Fashion Brands show the product colours in this way [11]. It is possible to conclude that Luxury Fashion Brands mostly opt for separate pages for the same product in different colours. It is possible to observe that the Portuguese LFDB prefer the change by click -on in the colour sample, whilst the International Luxury Fashion Brands opt

for change by scrolling down. Finally, none of the national and international brands studied presents colours in different products, automatic colour change, and a cursor swipe on their Website.

The macro layout variable is essential for the perception of the wearability of the pieces, it is fundamental so that customers can understand, through the layout of the product, how it will wear itself. The Portuguese LFDBs dispose of their products through catalogue style model, flat, with filling and model, leaving aside the disposition through a digital model, parcelled mannequin, mannequin, and hanging. On the other hand, International Luxury Fashion Brands dispose of their products through a model, flat, catalogue style, hanging, filled, parcelled mannequin and digital model, leaving aside the disposition through mannequin.

Finally, the macro-matching variable is also very important for the Portuguese LFDB online shop because it allows the customer to quickly identify, despite the inexistence of the assistant (as found in the physical shop), through a click similar or complementary pieces to what they are looking for. In addition to assisting the customer, it is also a way of increasing online shop sales. In the matching variable, it can be observed that only half of the Portuguese LFDB use this functionality and only present suggestions for each product, whilst the International Luxury Fashion Brands, besides the recommendations for each product, also present suggestions for look/season. None of the national and international brands makes the interactive combination available on their Website.

4.3 Results from the Interviews

According to the interview participants, being present online is an excellent way for Portuguese LFDBs to keep in touch with their customers and attract potential customers. As stated by one of the participants, Josefina's brand was born online "because in terms of management, for now, it makes sense, the team can be smaller (...), and it is online that represents a large part of sales". The other two interviewed brands, Carla Pontes and Katty Xiomara were created only in an offline environment, however, they chose to be present online. Carla Pontes said that she entered the digital market because she realised that "it would be a market of interest (...) and a future market". Katty Xiomara stated that she went digital due to some "positive experiences with aggregators, (...) although the demand is not significant", it seemed essential to have a more direct experience.

4.3.1 Current Digital Strategy

This category aims to determine the focal points of the digital strategies of Portuguese LFDB. The message they want to communicate online, how they distinguish themselves from other brands/designers, and if they create partnerships with other brands to increase their reach customers/potential customers.

According to what the interviewees explained, the current e-commerce strategy of Josefina's brand focusses mainly on PR, offers, partnerships, and constant product launch. On the other hand, Carla Pontes focusses more on paid advertising for other countries, intending to win new customers, and on communication and engagement with the brand's values to take the more human side of the brand to digital. Katty Xiomara is betting on greater fluidity of content and direct contact with her digital tribe.

LFDBs are also betting on communicating their message online. "We want to pass the message that we are not just a brand of shoes, but that we believe in values that are part of us, namely in women empowerment". (Josefinas). "The idea of non-temporality is very strong in our brand and also the idea of quality (...). The idea of pertinence in production, that is, in a non-mass production, we increasingly defend the idea of a conscious and ecological production". (Carla Pontes). Need for sustainability (...), it is not viable, neither for the company nor the planet, to produce pieces to archive. (...) We make the pieces that have better sales results and combine them with some limited-edition pieces. (...) Convincing the customer to buy less but to buy better. (Katty Xiomara).

Josefinas differentiates itself from the other Portuguese LFDB through product and storeytelling. In turn, Carla Pontes and Katty Xiomara differentiate themselves by their more personal side. Carla Pontes states that she tries to maintain proximity with her online customers to diminish the distance created online, as it is more personal than a virtual brand. Josefinas, Carla Pontes, and Katty Xiomara make partnerships with other brands, however, Katty Xiomara does not make partnerships to increase the reach of customers and potential customers. However, she says that indirectly partnerships can bring new customers to the brand.

4.3.2 Online Challenges

This category aims to determine the main challenges that Portuguese LFDBs face in the digital environment. One of the challenges encountered online is the entry of luxury brands with big budgets in the digital environment.

"Large brands, more or fewer luxury brands, have entered the digital world in force and have budgets that small brands like us do not have". (Josefinas) "The luxury brands and those that have big investors behind them will be able to overcome some of the difficulties that we feel today through artificial intelligence, visualisation with the camera itself, augmented reality, and visualising the piece almost dressed on itself instead of on the mannequin". (Carla Pontes).

The interviewee Josefinas adds that the increase in online competition has increased the cost of advertising in this channel. And that the digital market, despite being infinite, is becoming saturated by everything and everyone. Another challenge pointed out by Carla Pontes is the difficulty in making customers perceive the pieces online because there is no direct contact between the customer and the product. Katty Xiomara adds that a challenge she faces is deciphering the online data, as the market is increasingly difficult to predict.

4.3.3 Omnichannel

The designers of the interviewed brands consider it essential to be present both online and offline, as some barriers make it difficult to sell their products only online.

“The online by itself in fashion, a thing that depends so much, depends a lot on the experience of dressing, I think it is quite complicated, that is, it cannot live alone. (...) We have noticed some reluctance, a difficulty for the final consumer to believe in entering straight away, looking for the brand and being convinced only by online. (...) Customers try it on, they get the contacts and then, as they already know what size it is and how it fits, later they go to see if it is still online, and then they buy it. (...) But we feel that this first more direct contact in the physical shop is almost always necessary”. (Carla Pontes). “Normally, those who buy from us online already know our brand, our quality, the materials we use and even our fitting”. (Katty Xiomara). “We only sell online (...) we do pop-up stores”. (Josefinas).

As verified in a previous section of this study, most Portuguese LFDB do not provide an end-to-end experience to their customers because, although 71% of the brands/designers allow commercial transactions on their Website and use social commerce and only 50% provide e-commerce assistant and none has shoppable videos and virtual shop. However, brands/designers recognise the need to integrate the two channels because the fashion market depends a lot on the experience of dressing, making it very difficult to convince the customer only online, which still presents several weaknesses regarding the presentation of products. Usually, those who buy Portuguese Luxury Fashion Brands/Designers online already know the brand’s products. For example, Josefinas, a digital luxury brand, creates pop-up stores to make up for the lack of physical evidence.

4.3.4 The Luxury Atmosphere and Characteristics of Online Luxury Products/services

The interviews also revealed how Portuguese LFDB recreated the luxury atmosphere experienced in the physical shop online and how the characteristics of online luxury products/services are demonstrated.

The exclusivity of Josefina’s brand is achieved through par-to-par creation for each client. Carla Pontes has minimal productions. Katty Xiomara states that she works more on a made-to-measure and made-to-order basis. Physical evidence is one of the most challenging features to recreate online. Josefinas tries to make up for the lack of physical evidence through pop-up stores and facilitates the exchange and return. In turn, Carla Pontes uses closer photographs and videos that can show details. The designer says that when the brand presents the collections, it shows the storey behind each one and tries to evolve through backstage images. These videos conceptually involve the collection, and finally, synopsis texts to clarify the ideas passed previously. Katty Xiomara says they are still trying to discover and understand how to create an online luxury environment.

4.4 Final Notes on the Results

This article demonstrates how Portuguese LFDBs use their Website to create a luxury atmosphere and communicate and engage their current and potential customers. Online brand value is conveyed through the presentation of the brand storey on their Website (86%), product information (71%), email registration (64%), and corporate information (57%). Through the interviews, it was possible to conclude that brands/designers also communicate brand value through the craftsmanship and quality of their products, customer experience, packaging where they ship products, sending personalised cards and lookbooks, and delivering products within 24 h.

Portuguese LFDB do not yet use all the technologies and tools at their disposal to make up for the lack of physical evidence online. The products of the brands/designers are mainly presented through photographs with front, side, and back views, in which it is possible to zoom in or out. The products are arranged in a catalogue style model, with filling and flat. The different colours of the same product usually are presented on other pages, however, when they are shown on the same page, the colour change is either by scroll-down or click-on the colour sample. As such, brands/designers do not use 3D technology, augmented reality, 360° visualisation and multiple zooms, making it harder to transpose product features online.

The reason given by the participating brands in this study was that they do not use these technologies because they require significant investments. They choose to make up for the lack of physical evidence by facilitating exchanges and returns, closer photographs, presentation of the storey behind each collection, involvement through backstage images, videos of the collection that show details and synopsis texts.

Exclusivity is achieved through the availability of iconic products online, minimal productions, sometimes customer-specific, and direct or short distribution channels, not using community activation to deliver their products.

Most Portuguese LFDB market their products online and offline. Some brands/designers who sell their products only online occasionally create pop-up stores to buy them offline. Those that sell only offline make some of their products available on their Website.

5 Conclusions

This article provided important insights to explain the critical elements of LFDB digital strategies, particularly in what concerns Portuguese LFDB. This study demonstrates that it is essential that Portuguese LFDBs are present and market their products online, supported by an omnichannel strategy to ensure an effective relationship with their customers and satisfy their shopping experience [14]. They maintain the characteristics and differentiating elements of luxury products and services because the

online shopping experience should be equal to the shopping experience in one of the brand flagships so that the brand equity is not damaged [14].

Brands/designers face several online difficulties that affect the definition of their strategies for their Websites. Namely, small budgets directed to online, increasing competition and cost of online advertising, a saturation of the digital market, getting customers to perceive pieces through digital and understanding market volatilities.

As such, the Portuguese LFDBs define as critical elements of their e-commerce strategy: public relations, offers, partnerships with influencers, constant product launch, paid advertisements in other countries, communication and involvement of the brand's values, greater fluidity of content, more direct contact with the digital tribe, and creation of partnerships with other brands to reach potential customers. Allied to these elements, brands/designers communicate a message associated with the causes they support and bet on differentiation through the product, storytelling, and communication of the more human side of the brand. It is important to highlight those brands/designers still do not provide an end-to-end experience to their customers. However, they recognise the need to integrate the online and offline channels because the online depends heavily on the wearable experience. As such, brands/designers should continue to improve their omnichannel strategy so that their customers and potential customers can initiate contact with the brand online or offline. All offline touchpoints are also available online, such as the e-commerce assistant. In addition, brands/designers should present more corporate information and videos with interviews to know who is behind the brand. They need to improve the presentation and information about the products, accompanied by investment in 3D technology, augmented reality, and 360° visualisation to better understand the products through the Website.

5.1 Implications for Management

The study made it possible to identify some elements that managers of Portuguese Luxury Fashion Brands/Designers can use to improve their e-commerce strategies.

Based on the information collected and the conclusions drawn from the study, some of the contact points that brand/designer managers can establish with their customers along the consumer journey are suggested. In the attention phase, brands/designers work on customer acquisition, repeat purchase, online advertising, offers, public relations, sending press releases with the concept and photos, and creating ads on social networks. In the attraction phase, they use constant product launches. These campaigns emphasise their values, events that evolve the collection, links with other creative areas, syndication, special promotions/campaigns, and product launch promotions to test the product and discounts on newsletter subscriptions. The advice phase has the most significant weaknesses. Brands/designers only make their social networks and online communities available for customers and potential customers to talk and solve their doubts with other customers and with the brand. Brands/designers must invest in creating online communities. They are

a handy tool for their customers and potential customers to ask questions about the brand and its products and are part of a decisive stage of the advice phase of the consumer journey. In the action phase, brands/designers allow their products online and offline, however, most brands do not send emails to customers reminding them of the uncompleted purchase. It is important to highlight that email reminders should be included in the promotional emails that customers can reject if they wish to avoid feeling pressured. In the advocacy phase, brands/designers work EWOM using influencers to wear the brand and create events, offering personalised sales, after-sales service, conceptualise the collection and keep in touch with their customers by sending emails at the launch of new collections, invitations to events and newsletters with promotions only for current customers.

5.2 Limitations of the Study and Recommendations for Future Research

The main limitation of the present study is the sample size and not being representative of the population, however, data were found to provide important information about the digital strategies of Portuguese LFDB. It should also be noted that the data users of the performance of International Luxury Fashion Brands used for the comparative analysis with Portuguese LFDB had been collected three years earlier, so they may not reflect updates that have occurred in the meantime. On the other hand, it was not possible to interview all the brands/designers in the sample, which would allow a deeper understanding of their digital marketing strategies.

Regarding future research, it is recommended that this approach focussed on Websites be extended to other types of platforms in the future, namely social media and marketplaces. A greater focus on specific objectives, particularly conversion regarding e-commerce, would also be precious for assisting brands and managers working in the luxury fashion sector. Future studies could also include the perspective of customers and Internet users in the analysis. It is crucial to carry out studies continuously and systematically in this area, especially considering the rapid changes in the digital environment.

References

1. Achille, A., Marchessou, S., Remy, N.: The age of digital Darwinism: Enhance the customer experience and transform your business to survive and prosper in the luxury digital era. McKinsey & Company (2018). <https://www.mckinsey.com/industries/retail/our-insights/luxury-in-the-age-of-digital-darwinism>, last accessed 2020/12/14.
2. Bardin, L.: *Análise de Conteúdo*, vol. 70, 4ª ed. Edições, Lisboa (2018)
3. Chandon, J.L., Laurent, G., Valette-Florence, P.: Pursuing the concept of luxury: introduction to the JBR Special Issue on "Luxury Marketing from Tradition to Innovation". *J. Bus. Res.* **69**(1), 299–303 (2016)

4. Dubois, D.: Digital and social strategies for luxury brands. Thieme, W.M. (ed.) *Luxusmarkenmanagement*, pp. 327–337 (2017)
5. Hansen, R.: How fashion brands learned to click—a longitudinal study of online interactive and social media adoption by luxury fashion brands. **21** (2011)
6. Hansen, R., Bjørn-Andersen, N.: Cube assessment framework for B2C websites applied in a longitudinal study in the luxury fashion industry. *J. Theor. Appl. Electron. Commer. Res.* **8**(2), 3–4 (2013)
7. Hoang, L.: Has luxury and E-Commerce Finally Reached a Tipping Point? *Luxury Society*. <https://www.luxurysociety.com/en/articles/2020/11/has-luxury-and-e-commerce-finally-reached-tipping-point>, last accessed 2021/01/29.
8. Ko, E., Costello, J.P., Taylor, C.R.: What is a luxury brand? A new definition and review of the literature. *J. Bus. Res.* (2017)
9. Kotler, P., Kartajaya, H., Setiawan, I.: *Marketing 4.0: Moving from Traditional to Digital*. Wiley, Hoboken (2017)
10. Loureiro, S.M.C., Araújo, C.M.B.: Luxury values and experience as drivers for consumers to recommend and pay more. *J. Retail. Consum. Serv.* **21**(3) (2014)
11. Marques, P.: Análise das Estratégias de E-Commerce das Marcas de Moda de Luxo. Repositório P.Porto <http://hdl.handle.net/10400.22/10940>, last accessed 2021.01/21
12. Pantano, E., Servidio, R.: Modeling innovative points of sales through virtual and immersive technologies. *J. Retail. Consum. Serv.* **19**(3), 279–286 (2012)
13. Roux, E., Floch, J.M.: Gérer L'ingérable: La contradiction interne de toute maison de luxe. *Décisions Marketing* **9**, 15–23 (1996)
14. Rovai, S.: Digitalisation, luxury fashion and “Chineseness”: the influence of the Chinese context for luxury brands and the online luxury consumers experience. *J. Glob. Fashion Market.* **9**(2), 116–128 (2018)
15. Teixeira, S.F., Machado, P.L.: Venda Online de Marcas de Moda de Luxo: Contradição de conceitos ou desenvolvimento natural do negócio da moda de luxo? *Int. J. Mark. Commun. New Media*, Edição especial 4—Luxury Marketing, 122–147 (2018)
16. Wijaya, D.W., Fauzi, A., Arifin, Z.: The effect of online store atmosphere towards emotion and its impact on purchase decision. *Russ. J. Agric. Socio-Econ. Sci.* **64**(4), 82–93 (2017)
17. Yan, H., Qu, X.: Research on the marketing strategy of luxury goods. *Contem. Logistics* **5**(1), 42–46 (2014)

Correction to: Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution



Filomena Castro Lopes , Paula Morais , and Abílio Cardoso

Correction to:
Chapter “Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution” in:
J. L. Reis et al. (eds.), *Marketing and Smart Technologies, Smart Innovation, Systems and Technologies* 279,
https://doi.org/10.1007/978-981-16-9268-0_14

In the original version of the book, the following belated corrections have been incorporated in chapter “Critical Success Factors for BI Implementation in a Portuguese Higher Education Institution”:

The author’s name has been changed from “Abílio Cardoso Lopes” to “Abílio Cardoso” in the Frontmatter, the Backmatter and Chapter 14.

Moreover, the ORCID identifier provided for the author “Abílio Cardoso Lopes” was incorrect. Hence, this has now been removed, and the ORCID identifier for the author “Paula Morais” was missing, which has been updated as 0000-0002-0039-3532.

The updated version of this chapter can be found at
https://doi.org/10.1007/978-981-16-9268-0_14

Author Index

A

Abbasi, Maryam, 121, 129, 141, 365
Acosta-Rubio, Zamira, 551
Afonso, Carolina, 565
Agostinho, Vasco, 539
Almeida de, Hugo, 431
Alversia, Yeshika, 307
Arbaiza, Francisco, 595, 631
Aufreiter, David, 23

B

Baptista, Pedro, 121
Bâra, Gabriela-Ramona, 527
Barbosa, Belém, 717
Barta, Sergio, 575
Bem-Haja, Pedro, 431
Berend, Gábor, 23
Bessa, Maximino, 585
Bigné, E., 213
Blasco, María F., 233
Bogdanović, Zorica, 329
Bouwer, Anna, 37
Braga, Pedro, 497
Branco, Frederico, 351
Breitsohl, Jan, 641
Brito-Costa, Sónia, 431
Brito, João, 129
Brito, Pedro Quelhas, 497

C

Cachay-Marín, Claudia, 631
Cá, Frederico, 129
Calderón, Xavier, 269
Carbajal-Obando, Ninowska-Camila, 665

Cardoso, Abílio, 171
Cardoso, Tiago, 141
Carrasco, Ramón A., 233, 515
Carvalho, Aida, 485
Carvalho, João Pedro, 653
Casais, Beatriz, 259, 431
Cazorla-Artiles, José Manuel, 181
Chkoniya, Valentina, 419
Cicha, Karina, 471
Clara, Irina, 451, 485
Cobo, Manuel J., 233
Cóndor-Herrera, Omar, 443
Correia, Joana Sampaio, 617
Correia, Ricardo, 485
Côrte-Real, Ana, 419
Costa, Marcelo, 121
C.-Sánchez, Eleazar, 3
Cunha, Carlos Rompante, 463, 605

D

Despotović-Zrakić, Marijana, 329
Díaz-Padilla, Vidina Tais, 551
Duarte, Paulo, 539
Đurišić, Vladimir, 293

E

Esparza, Daisy Imbaquingo, 151
Espinoza-Nieves, Javier, 595
Estevão, Cristina, 539

F

Fátima Barros de, Maria, 79
Felizardo, Sónia, 431

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

J. L. Reis et al. (eds.), *Marketing and Smart Technologies*, Smart Innovation, Systems and Technologies 279, <https://doi.org/10.1007/978-981-16-9268-0>

Ferreira, Cátia, 717
 Ferreira, Miguel, 365
 Figueira, Álvaro, 79
 Flavián, Carlos, 575
 Foris, Diana, 527
 Fratrič, Alexandra, 23
 Fuentes-Medina, M. L., 213

G

Galán, José J., 515
 Gallardo-Echenique, Eliana, 631, 665
 Gonçalves, Gisela, 539
 Gonçalves, Ramiro, 351
 González, Karina Valarezo, 279
 González-Martel, Christian, 181
 Guaña-Moya, Javier, 269
 Guerra, M. J., 339
 Gurrea, Raquel, 575

H

Hán, Jan, 395
 Hipólito, Vanessa, 565
 Hirata, Sadayo, 93

I

Issá, Riaze Miguel Ribeiro, 191

J

Jorge, Filipa, 585
 Jose, Joemon, 641

K

Kaščelan, Ljiljana, 293
 Kato, Takumi, 69

L

López, Eduardo Parra-, 551
 López, María Judith, 389
 Labus, Aleksandra, 319
 Latorre, Antonio, 515
 Lavoura, Maria João, 79
 León, José Jácome, 151
 Lošťák, Michal, 395
 Loke, R. E., 243
 Lopes, Filomena Castro, 171
 Lopes, Luca, 365
 Lopes, Luisa, 605
 Losada, Nieves, 585

M

Macháčková, Karolina, 395
 Magalhães, Miguel, 693
 Mălăescu, Simona, 527
 Mang, Stefan, 23
 Marques, Catarina, 223
 Marques dos Santos, Jose Paulo, 191
 Marques, Luis, 365
 Martins, José, 351
 Martins, Pedro, 121, 129, 141, 365
 Melo, Miguel, 585
 Mena, Ricardo, 653
 Mendonça, Vítor, 605
 Mezarina, Melina, 665
 Mitrović, Svetlana, 319
 Monteiro, Francisco, 351
 Morais, Elisabete Paulo, 451, 463
 Morais, Paula, 171
 Morgado, Sónia M. A., 105
 Morini-Marrero, Sandra, 213
 Moro, Sérgio, 223
 Moutinho, Nuno, 259

N

Naranjo, David, 269
 Nešković, Stefan, 319
 Nogueira, Mafalda, 675

O

Ohme, Rafal, 419
 Oliveira, Luis, 129
 Ortega, MacArthur, 151

P

Paiva, Teresa, 451
 Penela, João, 141
 Popović, Snežana, 319
 Pratas, Joaquim, 703
 Prigita, Marleen, 307

Q

Quiña, Gissela, 151

R

Radenković, Miloš, 319
 Ramos-Galarza, Carlos, 383, 389, 443
 Reis, José Luís, 703, 717
 Reykowska, Dorota, 419
 Rivera, Mario Román, 279

Rodrigues, Margarida, [121](#)
Rogić, Sunčica, [293](#)
Romero-Domínguez, Laura, [3](#)
Rosa, André, [141](#)
Roxo, Mafalda Teles, [497](#)
Rutecka, Paulina, [471](#)

S

Salamea-Palacios, Christian, [269](#)
Saltos-Echeverria, Tatyana, [151](#)
Sánchez-Medina, Agustín J., [3](#)
Sanchez, Tarquino, [269](#)
Santos, Arlindo, [463](#)
Santos, Fernando Pinto, [693](#)
Silva, Bruno, [223](#)
Silva-Barragán, Micaela, [383](#)
Silva, F. O., [339](#)
Simões, Dora, [617](#)
Sista, Bruno, [259](#)
Skulimowski, Andrzej M. J., [55](#)
Sousa, Nuno, [585](#)
Stadlmann, Christian, [23](#)
Steentjes, J. A. A., [243](#)
Stojanović, Danijela, [329](#)
Sunthornpan, Saruda, [93](#)
Szántó, Zsolt, [23](#)

T

Teixeira, Mário Sérgio, [585](#)
Teixeira, Sandrina, [717](#)
Travar, Irena, [551](#)

U

Ueberwimmer, Margarethe, [23](#)

V

Varandas, Clésia, [675](#)
Vasconcelos, Maria João, [703](#)
Vaz, Bruno, [79](#)
Vidal, Juan, [233](#)
Volkmar, Gioia V., [15](#)

W

Wanzeller, Cristina, [121](#), [129](#), [141](#), [365](#)
Warke, Oliver, [641](#)

Y

Yordanova, Zornitsa, [407](#)

Z

Zelený, Jiří, [395](#)