Springer Series on Cultural Computing

Vladimir Geroimenko *Editor*

Augmented Reality and Artificial Intelligence

The Fusion of Advanced Technologies

Springer

Springer Series on Cultural Computing

Cultural Computing is an exciting, emerging field of Human Computer Interaction, which covers the cultural impact of computing and the technological influences and requirements for the support of cultural innovation. Using support technologies such as artificial intelligence, machine learning, location-based systems, mixed/virtual/augmented reality, cloud computing, pervasive technologies and human-data interaction, researchers can explore the differences across a variety of cultures and cultural production to provide the knowledge and skills necessary to overcome cultural issues and expand human creativity.

This series presents monographs, edited collections and advanced textbooks on the current research and knowledge of a broad range of topics including creativity support systems, creative computing, digital communities, the interactive arts, cultural heritage, digital culture and intercultural collaboration.

This Series is abstracted/indexed in Scopus.

Vladimir Geroimenko
Editor

Augmented Reality and Artificial Intelligence

The Fusion of Advanced Technologies

Springer

Editor
Vladimir Geroimenko
Department of Informatics and
Computer Science
The British University in Egypt
Cairo, Egypt

ISSN 2195-9056 ISSN 2195-9064 (electronic)
Springer Series on Cultural Computing
ISBN 978-3-031-27165-6 ISBN 978-3-031-27166-3 (eBook)
https://doi.org/10.1007/978-3-031-27166-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

With sincere gratitude to the Faculty of Informatics and Computer Science at The British University in Egypt.

Preface

This book is the first research monograph that explores a new research field and practical applications produced by combining two of the most advanced and powerful technologies available in today's world, namely artificial intelligence (AI) and augmented reality (AR).

The authors consider practical, theoretical, and cultural aspects of "AI-powered AR" and "AR-enriched AI", and their use in a large variety of areas, such as education, medicine, healthcare, dentistry, pharmacy, active lifestyle, smart services, fashion, retail, recommender systems, and several others.

The monograph is essential reading for researchers, practitioners, and technology developers, for students (graduates and undergraduates), and for anyone interested in building a comprehensive understanding of the emerging fields of "intelligent augmented environments" and "artificial intelligence presented through augmented reality".

The book is written by a team of 50 researchers and practitioners from 16 countries (Australia, Brazil, China, Egypt, France, Greece, Italy, Japan, Portugal, Romania, Serbia, Slovenia, Spain, UAE, Ukraine, and USA), which has enabled a thorough coverage of emerging and previously unexplored subject areas.

This book can be considered as part of a series of eight pioneering monographs published by Springer on the same subject of augmented reality and with the same editor, namely:

- *Augmented Reality Art: From an Emerging Technology to a Novel Creative Medium.* Geroimenko V (Ed), Springer, 2014
- *Augmented Reality Art: From an Emerging Technology to a Novel Creative Medium.* Geroimenko V (Ed), 2nd Edition, Revised and Updated, Springer, 2018
- *Augmented Reality Games I: Understanding the Phenomenon of Pokémon GO.* Geroimenko V (Ed), Springer, 2019
- *Augmented Reality Games II: The Gamification of Education, Medicine, and Art.* Geroimenko V (Ed.), Springer, 2019
- *Augmented Reality in Education: A New Technology for Teaching and Learning.* Geroimenko V (Ed), Springer, 2020

- *Augmented Reality in Tourism, Museums, and Heritage: A New Technology to Inform and Entertain.* Geroimenko V (Ed), Springer, 2021
- *Augmented Reality Art: From an Emerging Technology to a Novel Creative Medium.* Geroimenko V (Ed), 3rd Edition, Revised and Updated, Springer, 2022
- *Augmented Reality and Artificial Intelligence: The Fusion of Advanced Technologies.* Geroimenko V (Ed), Springer, 2023

The book's 19 chapters, which can be read in sequence or randomly, are arranged in four parts as follows:

Part I "Practical, Theoretical, and Cultural Aspects of Integrating Augmented Reality and Artificial Intelligence" includes five chapters (Chaps. 1–5).

Chapter 1 "Mixed Reality and Deep Learning: Augmenting Visual Information Using Generative Adversarial Networks" summarizes the theory of generative adversarial networks (GANs), overviews the evolution of GAN methods that have either been used in published augmented reality research or have a strong potential for future AR work, and presents interesting use cases of GANs in AR. GAN is one of the most innovative deep learning techniques. Since it was first introduced in 2014, many variations have been proposed. Recent implementations can create photo-realistic high-fidelity visual content that is virtually indiscernible from real-world images. In AR, photo-realism is crucial for the user to perceive the composed reality as belonging to the same world. The composition of virtual objects into the real-world background must be done in a plausible way, with semantically aware object placement, consistent illumination, and convincingly cast shadows. Various GAN architectures were proposed that address one or another of these challenges. The discussed methods are categorized into three distinct groups that correspond to unrelated bodies of research: GANs for image composition, AR face filters, and 3D model generators.

Chapter 2 "Augmented Reality User's Experience: AI-Based Data Collection, Processing and Analysis" presents an AI-based approach to the systemic collection of user experience data for further analysis. This is an important task because user feedback is essential in many use cases, such as serious games, tourist and museum applications, food recognition applications, and other software based on augmented reality. For AR game-based learning environments, user feedback can be provided as multimodal learning analytics (MMLA) which has been emerging in the past years as it exploits the fusion of sensors and data mining techniques. A wide range of sensors have been used by MMLA experiments, ranging from those collecting students' motoric (relating to muscular movement) and physiological (heart, brain, skin, etc.) behaviour, to those capturing social (proximity), situational, and environmental (location, noise) contexts in which learners are placed. Recent research achievements in this area have resulted in several techniques for gathering user experience data, including eye-movement tracking, mood tracking, facial expression recognition, etc. As a result of user's activity monitoring during AR-based software use, it is possible to obtain temporal multimodal data that requires rectifying, fusion, and analysis. These procedures can be based on artificial intelligence, fuzzy logic, algebraic systems of aggregates, and other approaches. This chapter covers theoretical and practical aspects of handling AR user's experience data, in particular, MMLA data. The chapter gives an overview of sensors, tools, and techniques for MMLA data gathering as well as presenting several approaches and methods for user experience data processing and analysis.

Chapter 3 "Digital Dreams: Speculative Futures for Artificial Intelligence, Augmented Reality, and the Human Brain" explores possible future interactions between artificial intelligence and augmented reality. Taking an interdisciplinary approach, the chapter considers dreaming in both humans and machines as a lens through which to investigate how future integrations of AI and AR may support both creativity and learning. It begins by exploring the creative potential for immersive world building with AI and AR, as well as its potential impact on human perception. It then examines how AR and AI could provide moments of revelatory learning for individuals by uncovering hidden data structures that influence our daily lives. Next, the authors consider possible negative impacts of combining AI and AR technology; how their vulnerabilities to bias, hacking and invasions of privacy, could be exploited for nightmarish ends. Finally, they offer some thoughts regarding speculative futures for AI and AR, as fields that compel on-going creative and critical engagement by artists, designers, and technologists.

Chapter 4 "Augmented Galatea for Physical Pygmalion: A Phenomenological Approach to Intimacy in *VTubers* in the East Asia Region" analyses the effects on the intimacy of introducing *VTubers* in the East Asia region from a phenomenological perspective. Digital technologies are pervasive, and they are becoming intimate in many different ways. There are dating apps dominated by algorithms and artificial intelligence generated to be digital companions of the users. Much attention has been given to the creation of virtual characters who can be controlled by artificial intelligence for their potential use in augmented reality and metaverse. China is one of the predominant countries implementing digital technologies within society in order to shape a novel way of living in the world. Chinese virtual characters are used by millions of people in their daily activities. For example, *VTubers* are virtual avatars created by an entertainment company to be used for social media. The virtual characters generated in augmented reality and usable in metaverse can have deep repercussions on who we are and how we live our intimacy by becoming intimate with us. This chapter focuses on the implication and the use of such virtual characters for developing intimacy with the users through phenomenology and post-phenomenology. The chapter is divided into three parts. The first part considers the introduction of *VTubers* into the Chinese cultural context and the impact these characters have on Chinese daily life (by relating the societal and Augmented Reality applications). The second part focuses on phenomenology to highlight the links binding people together and to show how people are shaped through this intimate connection and exposure to others. The third part intertwines *VTubers* in China with the findings highlighted in phenomenology to show how virtual characters can generate different subjectivities through their actions in daily contexts.

Chapter 5 "Augmenting Artificial Intelligences in Fiction: Evolving from Primordial Internet Memes to Cybergods of Disruption" begins with a comparison between an augmented reality-based comic book series entitled *Wintermoot*, where artificial intelligence software is used in portions of the books, and the first Internet memes produced by the Meme-Rider Media Team (an art collective founded in 1999 in Anchorage, Alaska), and then the chapter examines the usage of artificial intelligences in both the artistic sphere and in popular practice. The origin of the term "meme" begins in Richard Dawkins's *The Selfish Gene* from 1976, where he describes the equivalent of genes in culture, the "meme". Dawkins ascribes several biological equivalents to memes from genetic theory, leading to the ability of memes to congregate together at a critical mass large enough to create an artificial intelligence. Case studies look at the way *Wintermoot* employs artificial intelligence in augmented reality throughout the comic book series. From here, a history of the usage of artificial intelligences as a cybernetic phenomenon in literature is compared with the way artificial intelligence is used in reality. Pulling from the early history of meme theory in nonfiction before the advent of the Internet meme, this essay looks at the way memes propagate in society and how algorithmic thinking has produced social phenomenon beyond the sphere of the purely aesthetic and into a world of cybernetic networks. This is illustrated through the ways artificial intelligences are created (material negentropy), used, and understood in literature and cinema. Comparisons are also made between the advent of social practice and mirror world theory in the art world, as well as the way memes evolved on the Internet, concluding with a look at examples of artwork being produced in augmented reality that incorporates artificial intelligence as an artistic method for generating content, specifically the AIs in the *Wintermoot* series.

Part II "The Educational Use of Intelligent Augmented Environments" comprises five chapters (Chaps. 6–10).

Chapter 6 "Artificial Intelligence, Augmented Reality and Education" proposes several tools based on artificial intelligence and augmented reality to promote the transmission of knowledge in the educational context. AI is already used to provide assistance as part of educational processes (Squirrel AI, Google Classroom) and can generate effective methodological and assessment resources (Alexa, SIRI, NVIDIA Canvas, WriteToLearn, Lingvist, Duolingo, Gradescope, Smartick). In addition, the combination of AI and AR facilitates the learning of scientific, geographical, mathematical, and social content through the use of interactive scenarios (Metaverse, Munzee, Goosechase, Blippar, PlugXr or ActionBound, Bodyplanet, Arloon Geometry, Chromville Science, Explore the World). The objectives of this chapter are to analyse the functionality, advantages, and limitations of the use AI and AR in an educational environment, to illustrate its practical applicability and to create a table/guide with didactic examples.

Chapter 7 "Artificial Intelligence, Machine Learning and Extended Reality: Potential Problem Solvers for Higher Education Issues" investigates the potential of new digital technologies and innovative tools in solving higher education issues. Higher education establishments in both Serbia and Romania have experienced dramatic challenges under the process of building up a resilient educational system defined though sound cross-cutting competences according to market demands. Some changes resulted from the EU's transformational reforms approved, the new technologies applied, and teaching and learning methodologies implemented. The primary examples of technology are massive open online courses (MOOCs) and e-learning platforms, artificial intelligence, machine learning, extended reality, and blockchain. Examples of innovative teaching techniques include problem-solving, games, interactive instruction, the development of transversal competencies, and knowledge transfer. The targeted and well-structured approach of the agile management within higher education has helped to incorporate some of the educational market needs, outlining the potential to get rid of internal injustices as well. The chapter focuses on the survey method used to look at the opinions of students from the Information Technology School in Belgrade (Serbia) and the Spiru Haret University in Bucharest (Romania). Students took part in an exploratory survey to find out more about their knowledge of extended reality, machine learning, and artificial intelligence as well as their perspectives regarding the difficulties and potential uses of these tools.

Chapter 8 "Augmented Reality and Artificial Intelligence in Education: Toward Immersive Intelligent Tutoring Systems" explores how the convergence of AI and AR can impact the educational domain and lead to the development of more interactive and immersive intelligent tutoring systems (ITS). This chapter presents the technologies of AI, AR, and ITS and examines their use in educational settings. Then, it showcases the findings of recent related studies, discusses the potentials and merits that their combination engenders in educational contexts, and provides future research directions. Based on the findings, it can be said that the convergence of AI and AR can result in the development of interactive, immersive, engaging, and personalized learning experiences. These experiences can be applied in all educational levels and have the potential to take into account learners' psychological, cognitive, and motivational states as well as their unique characteristics, knowledge, preferences, interests, and performance. These experiences can help to meet the educational demand, positively impact teaching, and learning activities, and increase learning outcomes and motivation.

Chapter 9 "Augmented/Virtual Reality and Artificial Intelligence in Dental Education and Research" states that with the advent of "information technology" phenomenon, such as deep learning (DL), machine learning (ML), and artificial neural networks (ANN), dentistry has taken a new turn. Application of information technology in dental education and research, and its interaction with strategic frameworks, professionalism, and patient care is of paramount importance. The role of virtual reality (VR), augmented reality (AR), and artificial intelligence (AI) in improving dental school education and training has been primarily driven to overcome traditional teaching limitations. To overcome these short comings, several virtual reality systems have been adopted with an engagement of wide-ranging sensory channels by virtual simulation. This includes fully immersive VR (e.g. head-mounted displays, HMDs), semi-immersive VR (e.g. cave automatic virtual environment, CAVE), and non-immersive screen-based VR (computer-aided design systems, CAD). This chapter provides an insight of current trends and future forecasts of application of VR, AR, and AI in dentistry-related education and research.

Chapter 10 "The History of Furniture Objects: An Intelligent Augmented Reality Application" investigates how AR processes and user experience can be improved with machine learning (ML) techniques and proposes a solution for an intelligent AR mobile application named "FURNITURE". The fusion of AI with AR is not a new initiative but at the same time, its potential is yet to be maximized. Currently, artificial intelligence (AI) is applied to almost every domain. As is the case with other technologies, augmented reality can take advantage of the current advancements in AI to allow the creation of a new class of AR applications. The FURNITURE application was designed with two complementary functionalities: (a) to use an ML trained model to decide if an image of a piece of furniture corresponds to a pre-defined style, i.e. is authentic; (b) to augment the piece of the furniture shown in a real context, with a history of that furniture style, and a 3D model representative for that furniture style, the augmentation being conditioned by a previous validation of authenticity. The FURNITURE application uses Google AutoML framework to train a ML model for image classification, and Wikitude AR framework for the development of the mobile AR application. The designed purpose of the AR application is to support students in the fields of art, interior design, architecture and art history, professionals, and the public. The application demonstrates the concept and is open for future extensions, by expanding the object list that can be classified as corresponding to a certain style.

Part III "Augmented Reality and Artificial Intelligence in Medicine, Healthcare and Physical Activity" consists of five chapters (Chaps. 11–15).

Chapter 11 "Meta-Patients: Using Mixed Reality Patients and an AI Framework for Simulating Life-Like Clinical Examinations" addresses an augmented learning experience created for the Griffith University School of Nursing and Midwifery in 2022. An interdisciplinary research team from Griffith University (Australia) deployed the first iteration of an application for students in response to the difficulties imposed through the previous two years. The impacts of the COVID pandemic bought challenges to the Bachelor of Nursing program, particularly in relation to student competency in the physical assessment of patients, through objective structured clinical examinations. This pilot study introduced life-like, simulated patients, designed, and rendered within unreal engine to the students. The patients were accessible through cross-platform applications, including mixed reality devices. Students were also able to interact with patient information communicated using the AI framework afforded by Microsoft PowerApps packaged in a bespoke SharePoint site. Student participants were interviewed as part of the development process and approved of augmented and mixed reality as successful platforms for the deployment of the simulated patient scenarios within Microsoft Teams, and available through a mobile application and mixed reality device.

Chapter 12 "AI-Powered and 'Augmented' Dentistry: Applications, Implications and Limitations" deals with the possibilities that artificial intelligence and augmented reality provide in everyday dental practice and education, as well as the possibilities for the advancement of this technology in future. The implementation of augmented reality and virtual reality in AI-powered dental technologies is rapidly evolving into applicable solutions for clinical practice as well as bridging the gap between the digital and physical world. The development of AR is leading to the development of a new field in dentistry, which can be called "augmented dentistry" that allows the use of new dental technologies to improve various aspects of clinical practice, virtual treatment planning, and the evaluation and modification of virtual treatment outcomes.

Chapter 13 "Augmented Reality and Artificial Intelligence: Applications in Pharmacy" addresses Artificial Intelligence and Augmented Reality applications in pharmacy and pharmaceutical sciences. AI collects and analyses the data and mimics human cognitive functions to make intelligent decisions. AR allows a combination of real-world and computer-generated three-dimensional visualization that can be applied in education, training, and research. The development of AI and AR assists healthcare professionals in providing more personalized treatment options and planning for complex procedures to deliver better outcomes and quality of care. It can be utilized in pharmacy practice, including pharmacy education, pharmacogenomics, drug development, personalized medication, and treatment predictions. In addition, AR and AI can also help to create an interactive learning environment for pharmacy students to learn about real-world experiences from a classroom setting. Furthermore, the incorporation of AI and AR can help improve how healthcare providers utilize electronic health records (EHR) to interact with patients efficiently. With the help of these technologies, there is great potential for the future of pharmacy to be more advanced and efficient in delivering patient care.

Chapter 14 "Artificial Intelligence and Augmented Reality in Physical Activity: A Review of Systems and Devices" describes the existing AI devices aimed at PA monitoring and the most relevant AR software focusing on movement. It presents the latest findings, classified according to their accessibility and economic cost, and describes the main characteristics, the basic instructions for use, and the practical applications of each of them. Within the domain of physical activity (PA), artificial intelligence has been applied primarily in the form of smart wearable devices, with the aim of improving people's living standards. Through the use of wearables (Xiaomi Smart Band, Fitbit Apple Watch) and cameras (Alfa AI, Zenia, Peloton Guide, AndroVideo, Sportvu 2.0), AI can monitor the user's movements and physiological signals and can provide feedback about them. Augmented reality is now being employed to visualize, promote, and motivate new forms of human movement. Although there are very innovative initiatives aimed at promoting the practice of PA in a virtual environment, very little is currently known about these environments. The use of AR has given rise to a myriad of proposals in all fields of motor research and development, ranging from psychomotor processes in children (FitnessMeter, AR Runner, DribbleUp) to educational/sports PA practice (Lü interactive playground, BEAM interactive floor, HADO), PA and health (PostureScreen, Complete Anatomy 2022), high-performance sport (My Jump Lab, Capture.U,) and programs to promote PA in the elderly (Obie interactive projector, CyberCycle). There is software that uses both AI and AR, for example, Alfa AI, Capture.U, and CyberCycle. The use of these two technologies allows for the practice of PA in a new dimension and enables personalized, real-time monitoring of a user's activity through innovative exercises.

Chapter 15 "Exergames, Artificial Intelligence and Augmented Reality: Connections to Body and Sensorial Experiences" introduces augmented reality as one of the most prominent technologies promoting immersive and sensorial experiences. The research aims to describe how artificial intelligence has been used through game engines in complex feedback systems (Deep Learning), as well as customizing efficient models to engage and monitor human interaction in an immersive way in activities such as health and well-being. It was possible to observe that different fields such as military training, education, cultural exhibition, and many others are applying activities and developing many interactions mediated through AR. In addition, with the rise of wearables, it is clear there is a growing need to expand mechanisms that make the experience more realistic and interactive. One of the possible proposals would be the expansion of immersive environments with AR projection. Another growing area is entertainment, bringing new connections with virtual reality in the newest generation of exergames. The intention is to connect body movements with games through the sensorial immersion of AR, empowering new possibilities for the evolution of video games. This chapter presents a review of literature with the initial experiences involving AR and exergames through deep learning and to discuss the possible future of this emerging connection of sensoriality and body movements, supported by the AR tools and devices.

Part IV "Combining Augmented Reality and Artificial Intelligence to Enhance Services, Retail and Recommendations" includes four chapters (Chaps. 16–19).

Chapter 16 "FUSE: Towards AI-based FUture SErvices for Generating Augmented Reality Experiences" explores the use of artificial intelligence to semantically understand the scene, to decide when to provide augmented reality content or to take other actions, and how to generate the augmented content. The chapter presents a survey of novel machine learning techniques that have a capability to assist future AI-driven smart AR services and to enable their generalization to a wide range of scenarios making them ubiquitous, adaptable, and robust for real-world usage. The key challenge considered by the authors is how to generalize these techniques to unprepared arbitrary environments in which the scene objects are automatically detected, recognized, and segmented in an intelligent way that allows the system to annotate the scene automatically. The chapter begins with an exploration of various methods and techniques which could enable the creation of AR services at such a high level of generalization. It provides a review of a variety of machine learning algorithms for object detection, recognition, and image segmentation. Then, it analyses 3D scene creation and reconstruction, which includes various 3D scanning approaches as well as techniques for generating context aware text and image annotations. Finally, the chapter considers the newest methods of view synthesis using neural networks known as 3D neural rendering and finishes with a discussion about the current and future challenges of AI-based services for generating AR experiences.

Chapter 17 "Smart Extended Reality in the Metaverse-Tailing: The Rise of New Retail Landscape" aims to investigate the rise of Metaverse by considering the merge of interactive technologies (such as Augmented and Virtual reality and Artificial Intelligence) into a new retail landscape. The research flow of this chapter is guided by three theories able to identify and clarify how a smart extended reality in the metaverse-tailing can be considered as the rise of new retail environment. First, the affordance theory of technology will help to understand the possibility of obtaining value from specific technology, deriving from the metaverse world. Secondly, the lens of regulatory engagement theory is helpful to understand the positive engagement of people during their experience in a specific object or environment thanks to the exploitation of augmented and virtual reality in a metaverse world. In line with the regulatory engagement theory, this chapter aims to underline the possible way of involvement in the metaverse during a shopping journey and exploiting potentialities deriving from augmented reality, virtual reality, and artificial intelligence, such as interaction, immersion, inspiration, and satisfaction. Finally, the self-determination theory identifies the possible intrinsic or extrinsic motivation that leads people to experience retailing in the metaverse-tailing as a fusion of several advanced technologies.

Chapter 18 "Artificial Intelligence and Extended Reality in Luxury Fashion Retail: Analysis and Reflection" starts with a declaration that retailing service is facing a rapid evolution through the incorporation of extended reality (XR) technologies and artificial intelligence algorithms. Within this umbrella, one can consider all real-and-virtual combined environments and interactions generated by computer, including augmented reality (AR), mixed reality (MR), and virtual reality (VR). This chapter brings together research on the experience of luxury fashion consumer retail and AI and XR technologies to identify critical gaps and open avenues for future research. It aims to analyse the incorporation and the potential evolution of AI and XR technologies in luxury fashion retail. The chapter presents the main theoretical concepts and the luxury fashion consumption cycle and offers a discuss about the human and non-human rights, privacy, and ethics at three levels: mega, micro, and nano. Also, the chapter contributes with a discussion about the rights, privacy and ethics involving humans, hybrid humans, and non-humans.

Chapter 19 "The Use of Artificial Intelligence and Mixed Reality in Preventing Natural Disasters: Practical and Legal Issues" describes the role artificial intelligence and augmented reality play in preventing natural disasters such as pandemics, earthquakes, cyclones. Disaster relief agencies can process large volumes of fragmented and complex data with the help of AI. Consequently, it will be able to generate valuable information that can be acted upon more quickly. Despite its infancy, AR will soon become a key digital tool across many industries. Most companies are implementing technologies such as cognitive AI, big data, augmented reality, and cloud computing. In the recent pandemic situation, AR has been used for remote assistance, diagnostics, checklists, and training. Enhancing e-health experiences by combining AR and AI can also help resolve legal disputes. As part of its legislative process, the European Union (EU) has prioritized this awareness. Due to this, it is vital to emphasize that these new technologies during disaster emergency management are also supported by European legal standards.

Lastly, we hope that the reader will not judge the book's editor and contributors too hastily. We have accepted the challenge of researching the emerging field in which the fusion of the two cutting-edge technologies is currently taking place, and we have done our best to bring to you this comprehensive work on the combined use of artificial intelligence (AI) and augmented reality (AR). Please go ahead and read the monograph. We hope sincerely that you will enjoy it.

Cairo, Egypt Vladimir Geroimenko

Contents

Contributors

M. Ali Darendeliler Department of Orthodontics and Paediatric Dentistry, Faculty of Medicine and Health, School of Dentistry, The University of Sydney, Sydney, Australia

Ivan Allegranti School of Law, University of Camerino, Camerino, Italy

Frederic Andres Digital Content and Media Sciences Research Division, National Institute of Informatics, Tokyo, Japan

Gopi Battineni Medical Informatics, Clinical Research Center, University of Camerino, Camerino, Italy

Rob Burton Department of Nursing, Midwifery and Health, Northumbria University, Newcastle, UK

Federica Caboni Department of Management, Alma Mater Studiorum, University of Bologna, Bologna, Emilia-Romagna, Italy

Jenny Jiaying Chen Department of Philosophy, East China Normal University, Shanghai, China;
Art Director, Longlati Foundation, Shanghai, China

Vagner Ramos Dantas Psychopedagogy Department, Centre of Education, Federal University of Paraíba, João Pessoa, Brazil

Manuel J. De la Torre-Cruz Faculty of Educational Sciences, University of Jaén, Jaén, Spain

Daniel Della-Bosca Queensland College of Art, Griffith University, Gold Coast, Australia

Daniel G. Dinu Business Administration Doctoral School, Bucharest University of Economic Studies, Bucharest, Romania

Robert Ditcham Queensland College of Art, Griffith University, Gold Coast, Australia

Vanide Alves dos Santos Psychopedagogy Department, Centre of Education, Federal University of Paraíba, João Pessoa, Brazil

Mateus David Finco Psychopedagogy Department, Centre of Education, Federal University of Paraíba, João Pessoa, Brazil

Narayan H. Gandedkar Department of Orthodontics and Paediatric Dentistry, Faculty of Medicine and Health, School of Dentistry, The University of Sydney, Sydney, Australia

Roberto Garetto School of Law, University of Camerino, Camerino, Italy

Vladimir Geroimenko Faculty of Informatics and Computer Science, The British University in Egypt, Cairo, Egypt

Dragoş Gheorghiu Doctoral School, National University of Arts, Bucharest, Romania;
Instituto Terra e Memória, Mação, Portugal

Eileen Grafton School of Nursing and Midwifery, Griffith University, Logan, Australia

Gary Grant School of Pharmacy and Medicine, Griffith University, Gold Coast, Australia

Jessica Herrington Eccles Institute of Neuroscience, Australian National University, Canberra, ACT, Australia

Marius Hodea University of Theatre and Film, Bucharest, Romania

Louise Humphreys Technical Partners Health, Griffith University, Gold Coast, Australia

Milena P. Ilić Information Technology School ITS-Belgrade; LINK Group Belgrade; Faculty of Contemporary Arts Belgrade, University Business Academy in Novi Sad, Belgrade, Serbia

Matjaž Kljun Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia;
Faculty of Information Studies (FIŠ), Novo Mesto, Slovenia

Valentin Kuleto Information Technology School ITS-Belgrade; LINK Group Belgrade; Faculty of Contemporary Arts Belgrade, University Business Academy in Novi Sad, Belgrade, Serbia

Bertrand Laforge Laboratoire de Physique Nucléaire et des Hautes Energies (LPNHE), Sorbonne Université, Paris, France

Georgios Lampropoulos Department of Information and Electronic Engineering, International Hellenic University, Thessaloniki, Greece;
School of Humanities, Hellenic Open University, Patras, Greece

Nicola Liberati Department of Philosophy, Shanghai Jiaotong University, Shanghai, China

Sandra Maria Correia Loureiro Iscte-Instituto Universitário de Lisboa and Business Research Unit (BRU-IUL) and SOCIUS, Lisbon, Portugal

Teresa Martínez-Redecillas Group HUM-943: Physical Activity Applied to Education and Health, Faculty of Humanities and Educational Sciences, University of Jaen, Jaen, Spain

Larisa Mihoreanu Faculty of Administration and Public Management, Bucharest University of Economic Studies, Bucharest, Romania

Rasa Mladenovic Department for Dentistry, Faculty of Medical Sciences, University of Kragujevac, Kragujevac, Serbia

Mihaela Moţăianu National University of Arts, Bucharest, Romania

Dan Păun Faculty of Physical Education, Spiru Haret University, Bucharest, Romania

Andreas Pester Faculty of Informatics and Computer Science, The British University in Egypt, Cairo, Egypt

Lucia Pizzichini Department of Management, School of Economics "G. Fuà", Polytechnic University of Marche, Marche, Italy

Sabarinath Prasad Department of Orthodontics, Hamdan Bin Mohammed College of Dental Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

Klen Čopič Pucihar Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia;
Faculty of Information Studies (FIŠ), Novo Mesto, Slovenia

Don Roosan Faculty of Pharmacy, Department of Pharmacy Practice and Administration, Western University of Health Sciences, Pomona, CA, USA

Alberto Ruiz-Ariza Group HUM-943: Physical Activity Applied to Education and Health, Faculty of Humanities and Educational Sciences, University of Jaen, Jaen, Spain

Alba Rusillo-Magdaleno Group HUM-943: Physical Activity Applied to Education and Health, Faculty of Humanities and Educational Sciences, University of Jaen, Jaen, Spain

Nathan Shafer Anchorage School District, Independent artist, Łuk'ae Tse' Taas Comics and Shared Universe, Structured Learning Classrooms, Alaska, USA

Jose Luis Solas-Martínez Faculty of Educational Sciences, University of Jaén, Jaén, Spain

Yevgeniya Sulema Department of Computer Systems Software, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine

Sara Suárez-Manzano Group HUM-943: Physical Activity Applied to Education and Health, Faculty of Humanities and Educational Sciences, University of Jaen, Jaen, Spain

Denise Thwaites Centre for Creative and Cultural Research, University of Canberra, Canberra, ACT, Australia

Matthew Wong Department of Orthodontics and Paediatric Dentistry, Faculty of Medicine and Health, School of Dentistry, The University of Sydney, Sydney, Australia

Livia Ştefan Bucharest, Romania

Domen Šoberl Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia

Part I
Practical, Theoretical, and Cultural Aspects of Integrating Augmented Reality and Artificial Intelligence

Chapter 1
Mixed Reality and Deep Learning: Augmenting Visual Information Using Generative Adversarial Networks

Domen Šoberl

Abstract This chapter summarizes the theory of Generative Adversarial Networks (GANs), overviews the evolution of GAN methods that have either been used in published Augmented Reality (AR) research or have a strong potential for future AR work, and presents interesting use cases of GANs in AR. GAN is one of the most innovative deep learning techniques. Since it was first introduced in 2014, many variations have been proposed. Recent implementations can create photo-realistic high-fidelity visual content that is virtually indiscernible from real-world images. In AR, photo-realism is crucial for the user to perceive the composed reality as belonging to the same world. The composition of virtual objects into the real-world background must be done in a plausible way, with semantically-aware object placement, consistent illumination and convincingly cast shadows. Various GAN architectures were proposed that address one or another of these challenges. The discussed methods are categorized into three distinct groups that correspond to unrelated bodies of research: GANs for image composition, AR face filters, and 3D model generators.

1.1 Introduction

In Augmented Reality (AR) and Mixed Reality (MR), where virtual and real objects are combined seamlessly for a more immersive experience, physical and photo realism play an important part. Ideally, the user cannot differentiate between the virtual and the real and so perceives the composition of both worlds as if it belongs to the same space, which is very challenging to achieve (Pereira et al. 2021). The virtual content must be composed with the real-world background in a plausible way, which involves consistent object placement, illumination, and shading (Niu et al. 2021). The geometry of the virtual object must be consistent with the geometry of the real world. Object size, rotation, and position must be chosen to conform

D. Šoberl (✉)
Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia
e-mail: domen.soberl@famnit.upr.si

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_1

with real situations, e.g., a horse should not appear to be floating in mid-air or to be smaller than a dog. The light from the real world must illuminate the virtual objects the same as it illuminates the real objects, and the virtual objects must cast shadows consistently with the real objects. If the content is computer-generated, it should appear photo-realistic, which means that the user cannot distinguish between the visual appearance of a real and a generated object.

With recent advances in deep learning and generative algorithms, it became possible to approach many of the above challenges in a data-driven manner. Instead of designing complex and computationally exhaustive algorithms to reconstruct scene geometry and estimate the illumination, a model can be trained to perform the desired transformations. Numerous publicly accessible repositories can provide thousands of images and shapes for training. New content, which is unique in appearance, yet semantically meaningful, can be created from existing images or 3D shapes. The existing content can also be visually enhanced or partially reconstructed. Much of this is possible due to the advancement of Generative Adversarial Networks (GANs) (Goodfellow et al. 2014), which are gaining more and more interest, with hundreds of papers being published each year (Aldausari et al. 2022). This technology has now advanced to the point where it is virtually impossible to distinguish between real and synthesized images (Choi et al. 2020; Karras et al. 2021). The impressive results have attracted a certain level of interest within the AR community, where the demand for photo-realistic computer-generated content is high. To better understand the impact of GAN development on AR research, we are going to explore recent publications on such applications and determine the most prominent current trends of research as well as potential future directions.

After a brief overview of how GANs work, we will review the timeline of GAN evolution, which spans almost a decade since this architecture was first introduced. The motivation for such an overview is to understand better the different flavours of GAN, their potential, and their limitations, which may cast some light on why a particular architecture was chosen for one or the other AR application. We will then continue with actual use cases that were published in various scientific journals and conference proceedings.

This chapter groups the publications on the use of GANs in AR into three research areas. The first group contains research work on image composition, object placement, and shadow generation—these challenges are all closely linked to AR methodologies. The second group represents a large body of research on AR face filters, varying from manipulating facial expression, pose, and hair colour to artificial aging and applying virtual makeup. The third group is still an emerging area of research, with yet uncertain applicability but a promising potential for future development. It is concerned with generating point clouds of 3D objects as an alternative to manually modelling 3D shapes or capturing them through sensors or scanners.

There are individual attempts to use GAN architectures for specific AR applications that cannot be categorized into any of the three groups. These are individual attempts that target a particular challenge or a gimmick. They are too few to be considered an established or emerging body of research and are discussed in a separate section at the end of this chapter.

1.2 How Do GANs Work?

There are many flavours of GAN architectures, so we will focus on the elements they all have in common. What makes a deep learning architecture classify as a GAN is a composition of two competing neural networks, one called the *generator* and the other the *discriminator*, as shown in Fig. 1.1. The architecture aims to learn the distribution of the given training set in such a way, that the generator (G) learns to generate new samples that conform to the distribution of the training set, while the discriminator (D) learns to distinguish between the generated and the actual samples. A new sample is generated by mapping a vector z (typically chosen randomly) from the *latent space* to the domain of the training set. The latent space is a lower dimensional space that represents a compression of high-level concepts of the training domain. The generator, therefore, acts as a decoder of latent values into the realistic samples of the training domain. The decoder receives a sample, either a real or a generated one, and decides whether the sample is real or fake. The most common output format of the discriminator is the probability of the input sample being real.

The two networks are trained by competing in a minimax game, with the ultimate goal of reaching the Nash equilibrium (Ratliff et al. 2013). Discriminator D is trained to maximize the probability of assigning correct labels $D(x) = 1$, $D(G(z)) = 0$ to real samples x and fake samples $G(z)$, while the generator G is trained to minimize the probability of D to assign the correct label $D(G(z)) = 0$ to fake samples $G(z)$.

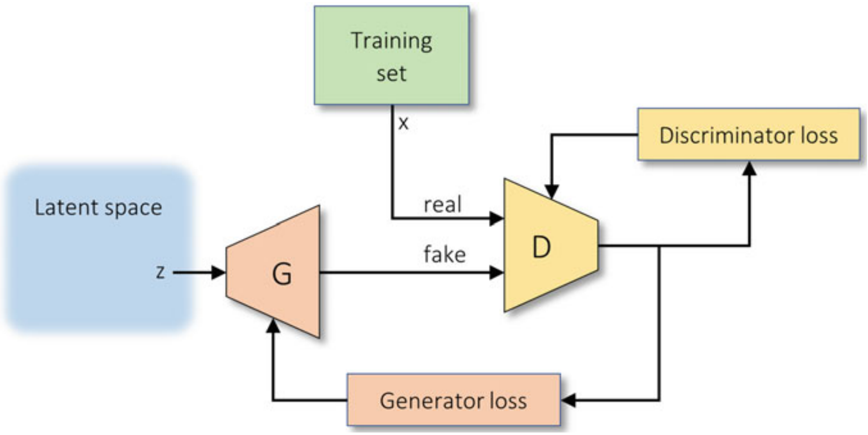


Fig. 1.1 A general architecture of Generative Adversarial Networks

Let m denote the size of the training batch. The generator loss function L_G and the discriminator loss function L_D are then defined as follows (Goodfellow et al. 2014):

$$L_G = \frac{1}{m}\sum_{i=1}^{m}\log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right),$$

$$L_D = \frac{1}{m}\sum_{i=1}^{m} -\left[\log D\left(x^{(i)}\right) + \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right].$$

The interpretation of the above equations is straightforward. The generator aims to minimize L_G, which means *fooling* the discriminator's output D to approach 1 with fake samples constructed from latent values $z^{(i)}$. The discriminator aims to minimize L_D, which means that the discriminator's output D should approach 1 with real samples $x^{(i)}$, and 0 with fake samples constructed from $z^{(i)}$. The adversarial setting is seen in the opposing aims for output D with fake samples. After several training steps, both networks ideally reach a point where they cannot further improve because the distribution of the generated samples matches the distribution of the real samples. The discriminator can no longer distinguish between the real and the generated samples, which means $D(x) = 1/2$ for any x.

1.3 Evolution of GANs for Image and Video Manipulation

Since GANs were first introduced in 2014, they have come a long way, with many variations and improvements being proposed. In less than a decade, their capabilities evolved from generating grainy low-resolution images to producing high-quality images and videos that are indistinguishable from real-world footage. Although GANs are not limited to generating only visual material, the latter is the most active domain of their research and probably the most important one when considering using GANs in AR and MR.

This section thoroughly reviews the evolution of GANs developed to generate visual content. All the GAN variants reviewed here are depicted on the timeline in Fig. 1.2. The timeline illustrates when a specific GAN was first introduced publicly (in many cases as a preprint), which may differ from its official publication in a journal or at a conference.

1.3.1 The Foundations

GAN framework was introduced in 2014 by Goodfellow et al. (2014). It quickly gained a lot of attention, primarily due to its ability to generate significantly better examples than existing generative methods of the time. Most notably, it overcame the

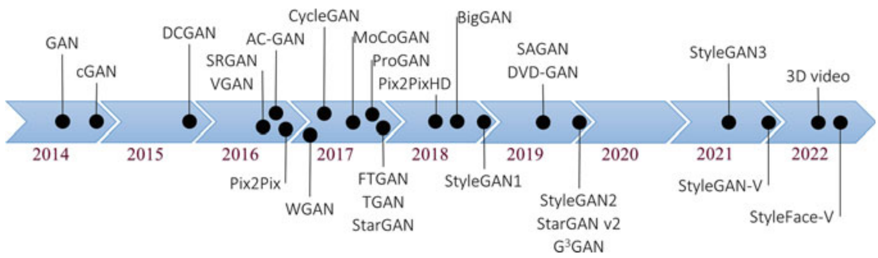


Fig. 1.2 Timeline of GAN architectures reviewed in this section

limitations of Variational Autoencoders (VAE) (Kingma and Welling 2014), which tend to produce blurry images. However, the original GAN framework has its own limitations, which numerous researchers addressed over time by introducing their own GAN variations. The research community often refers to the original GAN architecture as the *vanilla* GAN.

One of the main disadvantages of Goodfellow's architecture is that it uses *multi-layer perceptrons* (MLP)[1] in both the generator and the discriminator, which makes the training unstable. As a result, the generator can sometimes produce nonsensical outputs. Shortly after the introduction of the original GAN, Deep Convolutional GAN (DCGAN) (Radford et al. 2016) was proposed, which introduces the following changes: removing the fully connected hidden layers, using strided convolutional layers (O'Shea and Nash 2015) instead of fully connected layers, using the concept of batch normalization (Ioffe and Szegedy 2015), and using the Leaky ReLu (Maas et al. 2013) activation function.

1.3.2 Controlling the Process of Image Generation

The most notable downside of the vanilla GAN is the lack of control over the generated content. The generated samples could be of any category present in the training data, while some variations of the training data might not be represented in the output samples. Mirza and Osindero addressed this problem by introducing conditional GANs (cGANs) (Mirza and Osindero 2014). In a conditional GAN, both the generator and the discriminator are conditioned by additional information, such as class labels, text, or any other type of data, which helps generate images of a specific class. All training data must be labelled, which is often not the case when dealing with large image databases.

Auxiliary Classifier GAN (AC-GAN) was proposed to overcome this limitation (Odena 2016; Odena et al. 2017). Unlike cGAN, AC-GAN feeds the class labels only to the generator, while the discriminator must learn them. The discriminator, therefore, outputs two probabilities: the probability of the image being true/fake

[1] Multilayer perceptron refers to a neural network architecture with full connectivity between layers.

and the probability of the class. The AC-GAN model can perform semi-supervised learning by ignoring the class-label loss when a label is not given with the training image.

InfoGAN (Information Maximizing GAN) (Chen et al. 2016) was another slightly different approach to controlling the generated data. Unlike with cGAN, the auxiliary information that controls the generation process is unknown in InfoGAN. While training, InfoGAN tries to maximize the mutual information between the auxiliary value and the generated samples in an unsupervised manner. This way, it can learn meaningful features, such as pose, rotation, elevation, lighting, etc. In the case of human faces, it can learn the shape of a face, hairstyle, emotion, and presence of glasses. These learned features can then be used to control the generation process.

1.3.3 Translating Images to Images

Combining Deep Convolutional (DCGAN) and Conditional (cGAN) approach, a program called *pix2pix* was created, which introduced the concept of *image-to-image* translation, defined by the authors as translating one possible representation of a scene into another (Isola et al. 2017). Given a sufficiently large training dataset, the *pix2pix* program could translate sketches of faces into realistically looking portraits, outlines of cats or shoes into photographs of these objects, and architectural labels to photographs of facades, among others. The training set for this type of image-to-image translation must be paired, e.g., a sketch of a face with the corresponding portrait. Such pairing corresponds with the principle of supervised learning, where all training instances are labelled.

However, such labeled image datasets are scarce or unavailable in many cases. Take, for example, the challenge of translating a photograph into a painting as if drawn by Van Gogh. Obviously, no pairings exist of Van Gogh's paintings with photographs of the scenes that he painted, making the *pix2pix* approach inadequate for such a task. An unpaired image-to-image approach called CycleGAN (Zhu et al. 2017) addressed this problem by introducing a cyclic GAN architecture, which is composed of two generators and two discriminators. One generator-discriminator adversarial pair learns the mapping from image domain A to image domain B, while the other pair learns the mapping from B to A. CycleGAN aims to preserve cycle consistency, i.e., mapping an image from A to B and back to A (or vice-versa) should give us back the original image, which allows training to be done on unpaired sets of images. This method was successfully used for image style transfer, object transfiguration, season transfer, and photo enhancement.

To overcome the limitation of CycleGAN to learn the translation only between two domains, StarGAN (Choi et al. 2018) altered the CycleGAN architecture to allow simultaneous training on datasets from different domains. They achieved this by conditioning the generator with target domain labels and implementing the discriminator as a Wasserstein GAN, which we discuss in the next section. One of the limitations of StarGAN was that it learned deterministic mapping per domain, which means that it always produced the same output for a given source image and a domain label. StarGAN v2 (Choi et al. 2020) replaced domain labels with style code learned during training. Using these codes instead of labels enabled the generator to synthesize diverse images over multiple domains.

1.3.4 From Classification to Regression

A GAN discriminator is typically defined as a classifier, discerning between true and fake samples. A perfect discriminator would output $p(x) = 1$ when the input sample x is true, and $p(x) = 0$ when it is fake. This brings about a well-known problem called *the vanishing gradient problem*, which occurs when the generator's training falls behind the discriminator's training. As the discriminator is doing a better job than the generator, the gradient of the discriminator's loss function drops close to zero, leaving the generator with no gradient to update its loss. Wasserstein GAN (WGAN) (Arjovsky et al. 2017) addresses this problem by redefining the discriminator's task as a regression problem. The authors propose using the Wasserstein distance (also known as the *earth mover's distance*) as the distribution distance between real and fake samples. When the discriminator is trained to completion, it simply provides a perfect loss function to the generator. WGAN also turned out to be resilient to the problem of *mode collapse*, where the generator gets stuck in a small space of the real-data distribution and always produces the same output with little variety. A later variant of WGAN called WGAN-GP (Gulrajani et al. 2017) incorporated *gradient penalty* to even further improve training stability.

1.3.5 Higher Resolutions

Up to about this point in the history of GAN evolution, the outputs produced were low-resolution images, typically up to 128×128, or in some cases, 256×256 pixels. Before attempting to generate higher-resolution images, GANs were first considered as an upscaling tool. Super-Resolution GAN (SRGAN) (Ledig et al. 2017) was able to upscale low-resolution images by up to 4-times while preserving photo-realism. Such architecture is trained by downscaling high-resolution images and using the low/high-resolution pairs to learn a translation.

One of the first attempts to generate high-resolution images from scratch was the progressive GAN (ProGAN) (Karras et al. 2018), which implements a technique of incrementally adding convolutional layers to scale up to the desired resolution. The training starts with the 4×4 pixels architecture. When the first training iteration is done, the architecture is scaled to 8×8 pixels for the next training iteration, and so on, every additional layer doubling the output resolution. The authors went up to 1024×1024 pixels. A similar *coarse-to-fine* approach was taken by Pix2PixHD (Wang et al. 2018a).

It was soon demonstrated that high-resolution images could be generated without the need for explicit multiscale methods by increasing the number of parameters and scaling up the batch size. The proposed method, called BigGAN (Brock et al. 2019), was able to generate images of resolution up to 512 × 512 pixels and allowed fine control of the trade-off between the variety and the fidelity of the produced images. This was achieved by employing the SAGAN architecture (Zhang et al. 2019), which complements convolutional layers with *self-attention* modules that help coordinate the generation of fine details across different portions of the image. BigGAN used up to eight times larger training batches than existing methods, so it significantly outperformed them in the quality of generated images. It did, however, suffer from the problem of mode collapse, which was thoroughly investigated in the paper.

A few months after the publication of the BigGAN paper, the authors of ProGAN introduced the first version of the style-based GAN (StyleGAN1) (Karras et al. 2019), which soon gained much attention, especially for generating human faces. The architecture of StyleGAN was motivated by the style transfer techniques, which allow unsupervised separation of high-level attributes (such as pose or identity) from stochastic variations (such as hair, freckles, and skin pores). This, in turn, enables intuitive regulation of style mixing: for example, high-level aspects of the output images are generated from one source of images, while their stochastic variations are generated from another. The output images were of high fidelity in resolution up to 1024 × 1024 pixels. StyleGAN was used to build the well-known web page *This person does not exist*,[2] which generates a face of a non-existent human with every page refresh.

StyleGAN2 (Karras et al. 2020) fixed several quality issues of StyleGAN1, the most notable being the appearance of blob-like objects on the output images, which the generator created to circumvent certain design flows in the architecture. With the improved architecture, the training speed of StyleGAN2 was also up to 40% faster compared to StyleGAN1. The authors later observed that the image synthesis process depends on absolute pixel location, which results in the so-called *texture sticking* problem. As a result, the details appear to be glued on image coordinates instead on the surfaces of objects. The negative effects are clearly visible in interpolation videos,[3] where the illusion of solid, coherent objects moving in space is broken. StyleGAN3 (Karras et al. 2021) solves this problem by changing the generator in such a way that all signals in the network are interpreted as continuous. The results helped pave the way for generative models better suited to video and animation.

[2] https://this-person-does-not-exist.com/.

[3] The videos can be seen on the official project page https://nvlabs.github.io/stylegan3.

1.3.6 Video Synthesis

Following the success of GANs in generating images, there were several attempts to use GANs to generate videos (since a video consists of images with an additional time dimension). The first known GAN method to generate videos was called VGAN (Vondrick et al. 2016). The authors introduced a two-stream generative model—one stream used 2D convolutional layers to generate a static background, while the other used 3D convolutional layers (one additional dimension to process the temporal properties of a video) to generate a moving foreground. Both outputs were then combined into a video and fed into a discriminator. The generated videos were of low fidelity, containing 32 frames at 64×64 pixel resolution. Flow-and-Texture GAN (FTGAN) (Ohnishi et al. 2018) followed the same approach as VGAN but additionally separated the generation of optical flow from the generation of textures to produce videos that have more realistic motion. The fidelity of the videos was the same as that of VGAN.

Other research went in the direction of exploring the temporal properties of video generation. Temporal GAN (TGAN) (Saito et al. 2017) added a separate temporal generator to generate a set of latent variables, each corresponding to a frame in the video before these variables are transformed into images by the image generator. Motion-and-Content GAN (MoCoGAN) (Tulyakov et al. 2018) was able to learn a separation between motion and content, e.g., a separation between a person's identity and their facial expression, thus making it possible to generate videos of changing facial expressions. G^3GAN (Wang et al. 2020) further improved the process of separating motion and content by adding the third stream that modeled spatio-temporal consistency.

A quest for high-fidelity videos began with DVD-GAN (Clark et al. 2019), which was built upon the BigGAN architecture and could generate videos at resolutions up to 256×256 pixels and lasting up to 48 frames. However, real progress in this direction was made after the introduction of StyleGAN. Both MoCoGAN-HD (Tian et al. 2021) and StyleGAN-V (Skorokhodov et al. 2022) are built on top of StyleGAN2, which enables them to generate videos of resolution up to 1024×1024 pixels. However, MoCoGAN fails to generate videos of arbitrary lengths, while StyleGAN-V can generate videos of arbitrary lengths and framerate.

It is difficult to predict future research trends in video synthesis using GANs, but some recent publications indicate a few possible directions. Consistency and diversity of longer synthesized videos are still an open problem. Tim Brooks and his colleagues pointed out that StyleGAN-V tends to repeat the same content over time, e.g., a horse fails to move past an obstacle and morphs back and forth within a short window of motion (Brooks et al. 2022). They propose prioritizing the time axis and training the models on long videos at low resolutions and on short videos at high resolutions. They encourage further research on long-term changes in generated videos. Another trend may go in the direction of domain specialization. For example, StyleFaceV (Qiu et al. 2022) uses a pre-trained StyleGAN3 model to generate facial expressions in the synthesized videos that are not restricted to the domain of the training set. Recently, the first GAN attempts were made to synthesize a 3D video through monocular video supervision (Bahmani et al. 2022).

1.4 Image Composition in AR Applications Using GAN

Augmented reality is often faced with the challenging problem of composing a real-world background image with computer-generated foreground graphics in a plausible way. This means that the virtual object inserted into the real-world environment must have a reasonable scale, position, and lightning consistency. These problems were traditionally approached by designing explicit rules for finding appropriate positions or by inverse rendering techniques to estimate the geometry of objects. Although these approaches may be effective in some cases, designing them for diverse and complex settings is very challenging. Hence, various deep learning methods appeared in recent years to tackle some of these problems in an automated manner (Niu et al. 2021). Here, we review GAN approaches to image composition that tackle the challenges of object placement, image harmonization, and shadow generation.

1.4.1 Object Placement

The first known GAN approach to image composition is called Spatial Transformer GAN (ST-GAN) (Lin et al. 2018), which utilizes a Spatial Transformer Network (STN) (Jaderberg et al. 2015) for its generator. The generator warps (transforms) the foreground image before composing it with the background image. The discriminator then assesses the composite image, which computes the Wasserstein (WGAN) loss to train the generator. The main difference between ST-GAN and a conventional GAN is that ST-GAN learns to generate warp parameters instead of images as pixel arrays. The authors evaluated their method on the problems of placing furniture into indoor scenes and placing glasses on photos of celebrities (as in Fig. 1.3). The results showed that the augmented objects progressively moved towards more realistic positions with each update of the warping parameters. However, there are cases where ST-GAN fails to converge. A regularization technique was later proposed (Kikuchi et al. 2019) to guide warping parameters based on user-defined trust regions.

While ST-GAN learns only geometric transformations of augmented objects, Spatial Fusion GAN (SF-GAN) (Zhan et al. 2019) adapts the CycleGAN architecture to achieve realism in both the placement and the appearance. Therefore, SF-GAN uses two generators—the *geometry synthesizer*, similar to the ST-GAN generator, and the *appearance synthesizer*, inspired by the image-to-image translation GAN methods. The evaluation by the authors showed considerable improvements in SF-GAN image composition compared to ST-GAN.

Fig. 1.3 Correcting the position of glasses with ST-GAN.[4] Left: a headshot photo of a person. Middle: the headshot photo combined with an image of glasses. Right: ST-GAN readjusts the glasses to match the face

One of the limitations of ST-GAN is that it works only with fixed backgrounds and, therefore, cannot capture potential interaction between objects in the real world. For example, a chair and a table would form a composition in the 3D world, which would then be captured by a 2D projection. Such a composition would inevitably contain some level of object occlusion, which ST-GAN cannot generate. This problem can be approached using a self-consistent Composition-By-Decomposition network (CompGAN) (Azadi et al. 2020), which can decompose an image of two composed objects back to the respective individual object images. Instead of simply blending a virtual object against a fixed background, such a network can consider two objects in interaction. Moreover, such architecture can be trained on unpaired image datasets by using decomposition as a supervisory signal to train composition while minimizing the self-consistency loss. There are apparent advantages of being able to train the generator on unpaired data, e.g., when composing images of tables and chairs, we may not have an image of a particular chair-table composition, but we might have images of other chairs and tables together.

In image composition, *where* and *what* can sometimes be regarded as two separate questions. Lee and his colleagues proposed a two-step process, where the *where* and the *what* generative modules separately learn the probabilistic distributions of *locations* and *shapes* (Lee et al. 2018). The *where* generator applies an affine transformation over a unit bounding box and outputs the candidate regions for the new object. Given a semantic map, the *what* generator then generates plausible shapes and poses. In some applications, the *where* question may be the only concern for image composition, e.g., placing pedestrians in plausible locations on the background image of a street. To predict diverse locations of commonsense locations, a model called PlaceNet was proposed (Zhang et al. 2020), which can be trained on sparse real-world observations. More recently, Structure-Aware Image Composition (SAC-GAN) (Zhou et al. 2021) outperformed ST-GAN and CompGAN in generating

[4] The headshot photo is the courtesy of Marina Paldauf. Used with permission. The glasses were positioned and readjusted using the source code from the official ST-GAN repository at https://github.com/chenhsuanlin/spatial-transformer-GAN.

better object scales and PlaceNet in creating better semantic-aware locations. Moreover, SAC-GAN achieved stronger training supervision than these three methods by cropping the objects in the ground truth images. The authors demonstrated their image composition approach to the problem of inserting vehicles into cityscapes.

1.4.2 Harmonization and Shadow Generation

In some AR applications, simply placing a virtual object over a real-world background may suffice, e.g., mobile applications that put labels, icons, or other overlays over the camera stream. However, as we move towards immersive experiences of Mixed Reality, we expect virtual objects to appear as belonging to the real world, with consistent colours, lighting, and shadows. The process of achieving this is called *image* harmonization. Using GANs, image harmonization can be approached as an image-to-image translation problem, where the style of the background image is transferred to the foreground image (Zhan et al. 2020). However, generating shadows is somewhat more complex than ordinary image-to-image translation and a focus of much GAN-based research in image composition. Traditional methods of generating shadows require explicit information on illumination, material properties, and geometry, which is usually unavailable and has to be estimated. Such estimations are difficult to compute and give implausible results when inaccurate; hence a GAN approach seems very appealing.

The first known GAN-based approach to shadow generation was actually a shadow removal method called Mask-ShadowGAN (Hu et al. 2019), which used a cycle-consistent architecture similar to CycleGAN to train the model on unpaired images. An input shadow image is first translated to a shadow-free image. The subtraction of the two images results in a shadow mask that depicts the removed shadow's position and shape. The shadow-free image can then be composed with the shadow mask to retrieve the original image, thus preserving the cycle consistency. Such architecture can generate shadows on shadow-free images by providing a shadow mask, but this is not always practical.

The first known GAN-based shadow generator, ShadowGAN (Zhang et al. 2019), also requires an input mask and a shadow-free input image. However, the mask depicts the object, not the shadow. The shadow is then generated. ShadowGAN is trained on paired images. A training instance contains one or more real-world objects with a visible shadow and one virtually inserted object without a shadow, a mask of the inserted object, and the ground truth image with an appropriately rendered shadow of the inserted object. This method works with only one object being inserted into the image but can be applied iteratively for multiple objects.

ARShadowGAN (Liu et al. 2020) offers somewhat better results than ShadowGAN, taking the same input objects as the latter. The improved results come from a different architectural approach, which explicitly accounts for the real-world occluders using a dedicated *attention mechanism*. This mechanism takes as input the image and the object mask. It then outputs the feature map of real-world

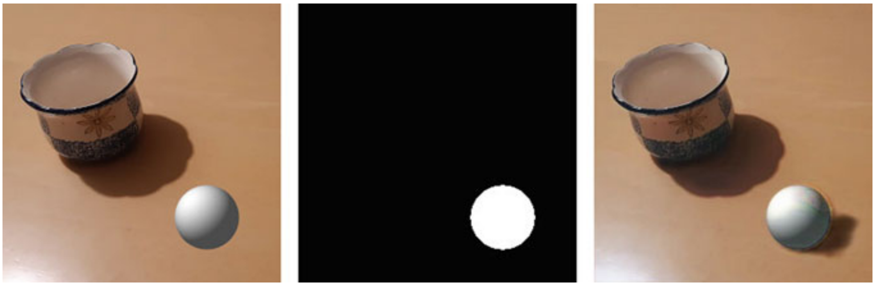


Fig. 1.4 Generating a shadow with ARShadowGAN.[5] Left: the input image contains a real-world background and a computer-generated ball in the foreground. Middle: the mask that depicts the foreground object. Right: the output image with the synthesized shadow

occluders and the feature map of their real-world shadows. These two feature maps are then used to synthesize a shadow of the inserted virtual object (see Fig. 1.4). The authors of ARShadowGAN also constructed a publicly available large-scale dataset of nearly 3000 instances to train and test ARShadowGAN models. The background images within the dataset are real-world photographs, and the inserted objects are all computer-generated. The shadows of virtual objects on the ground truth images are 3D rendered.

More recently, a method called SGRNet (Shadow Generation Network) was proposed (Hong et al. 2022) as a two-stage process for generating shadows—the shadow prediction stage and the shadow filling stage. Taking the same input type as ShadowGAN and ARShadowGAN, the shadow prediction stage first generates a binary shadow mask of the foreground object. The input image and the generated shadow mask are then fed through the shadow-filling stage, where appropriate illumination is applied to the masked region. This way, the architecture considers shadow shape and intensity separately. The evaluation done by the authors shows considerable improvement in generating realistic shadows compared to previous methods. In addition, the authors provide their own publicly available dataset DESOBA (DEshadowed Shadow-Object Association), which contains over 1000 images. Unlike the ARShadowGAN repository, where inserted objects are all computer generated, DESOBA contains only real-world images, where the shadows on the de-shadowed objects were removed manually.

1.5 AR Face Filters

Augmented Reality Filters (ARF) are computer-generated effects that are layered over a real-world video to enhance or augment the contents of the video. They have become trendy on social media platforms, especially the ARFs that manipulate

[5] The output image was generated using the source code from the official ARShadowGAN repository at https://github.com/ldq9526/ARShadowGAN.

human faces. Some researchers have explored the possibilities of using various GAN architectures to learn such filters from the vast amounts of publicly available face image repositories, such as *Labeled Faces in the Wild Home* (LFW) and *Large-scale CelebFaces Attributes* (CelebA). Besides the general question of how to change certain facial features or expressions, much attention has also been given to face aging and makeup filters.

1.5.1 Manipulating Facial Features

Conditional GANs (CGANs) offer controllability over the generated content. For example, one could construct a GAN to generate faces by controlling certain classes of appearance, e.g., gender, hair color, presence of glasses, etc. But the main challenge the researchers are facing is preserving the generated face's identity while changing such attributes. One of the first to address this problem was Kaneko with his colleagues, who proposed Conditional Filtered GAN (CFGAN) (Kaneko et al. 2017). They focus on how to construct a latent space that is disentangled, expressive, and controllable. A disentangled latent space allows the user to change a chosen attribute (e.g., hair colour) while preserving other visual features. An expressive latent space provides a wide diversity of attributes to choose from, i.e., the dimensionality of the conditional variable. The level of controllability can be associated with the granularity of the control variable, where continuous control offers the highest level of controllability. For example, the user could move a slider within the range between *young* and *old* to fine-tune the perceived age of the generated face. A smooth slider would represent the highest level of controllability for this attribute. The control variables in CFGAN are learned in an unsupervised manner. They can be determined intuitively after the training has been finished by trying out what features the control values affect.

Another approach towards face generation control is incorporating geometry awareness into GAN architecture. Geometry-Aware GAN (GAGAN) (Kossaifi et al. 2017) was the first GAN variant able to produce realistic images of faces conditioned by an input geometric shape. The geometry of the face is determined by a set of fiducial points—points that mark different facial components, e.g., 'left eye center', 'nose top', 'chin', etc. Once trained, GAGAN can control appearance and shape parameters separately (similar to CGAN). The identity of the face can therefore be preserved by fixing the appearance parameters, while the pose, morphology, and facial expression is manipulated by changing the shape parameters. The authors demonstrated the generality of their approach by also applying GAGAN to the faces of cats.

However, controlling facial expressions with fiducial points is not very intuitive. GANimation (Pumarola et al. 2020) attempts to control the generation of facial expressions through so-called Action Units (AUs) as a part of the Facial Action Coding System (FACS) (Ekman and Friesen 1978). Each AU is anatomically related to contractions of specific muscles so that each human emotion can be encoded as a vector of AU activations or magnitudes. A continuous change of the AU vector can then continuously animate a human expression, e.g., from a serious face to a smiling face. Unlike fiducial points, AUs are not person specific, meaning it should be possible to extract a facial expression from one image and transfer it to another while preserving the facial identity. Soumya Tripathy and his colleagues have thoroughly researched this problem (Tripathy et al. 2020, 2021, 2022). They proposed several approaches to neutralize or transfer the facial expression and the head pose from a *driving* image to a *source* image. Furthermore, they tackled the problem of preserving the source facial identity even when the driving image is of a different person and studied the motion prediction represented in the driving image.

1.5.2 Face Aging

A notable trend in GAN face manipulation research is facial aging. The traditional approach through physical-based modeling by simulating aging mechanisms of skin, muscles, and skull often requires large amounts of image sequences of the same person, which is extremely difficult to collect in practice. Generative approaches are typically prototype-based. This means that faces are first classified into age groups, then, the differences between average faces from each group are used to study the aging pattern.

Early GAN methods for face aging (Antipov et al. 2017; Wang et al. 2018b; Yang et al. 2018; Liu et al. 2019) classified faces into 4–6 non-overlapping age groups, where each group represented an age span of roughly 10 years. These approaches could, therefore, only account for long-term facial age synthesis. Nevertheless, they addressed some of the fundamental challenges of GAN-based aging, the most notable being preserving the identity of the generated face. Typically, the age group is used as a parameter for the conditional GAN to generate a face. The latent space must therefore be disentangled so that the person's identity can be preserved while switching the age condition. This approach requires an identity loss function that can identify a person regardless of age, which means that higher-level features, such as colors and textures (hair color, wrinkles, etc.), should not affect identity information (Wang et al. 2018b). Specific features can be made more important to improve the accuracy of age estimation, e.g., forehead and hair components (Yang et al. 2018). Personal attributes such as gender or race can also play an important role in facial aging, for example, a neural network could learn that a beard corresponds with older age groups and age a woman by generating a beard on her (Liu et al. 2019). An example of aging a person through a conditional GAN with one of the earlier methods is shown in Fig. 1.5. While some degree of aging effect was achieved, there was evidently still room for improvement.

More recent approaches (Sun et al. 2020; Huang et al. 2021) focus on the problem of short-term aging, which allows a more gradual transition from a younger to an older face. Instead of classifying the training instances into age groups, each instance is assigned an age distribution, meaning that an instance belongs to an age group

Fig. 1.5 Aging a person while preserving their identity. The leftmost image is the original one. The rest were generated[6] with the face-aging GAN proposed by Wang et al. (2018b)

with a certain probability. Each sample, therefore, contributes not only to learning its age group but also to learning the neighbouring groups. Consequently, these methods better capture the correlation between adjacent age groups. It goes along with the intuition that the aging process differs at different ages. The results of these newer methods seem much more impressive, but their reproducibility depends on the willingness of their authors to share the source code and the pre-trained models in the future.

1.5.3 Makeup Filters

A makeup filter is an image filter that applies virtual makeup to a real-world face image. Within the GAN community, makeup filtering is usually considered and defined as a style transfer problem, i.e., an image-to-image translation where a non-makeup image is translated to a makeup image or vice-versa. In most cases, the problem is approached in an unsupervised manner, in the sense that the training images are unpaired. BeautyGAN (Li et al. 2018) is the first known attempt to adopt a GAN architecture for makeup transfer. It adopts a cyclic GAN architecture with cycle consistency loss to employ an unpaired training dataset and the perceptual loss to preserve the identity between a non-makeup face and the generated makeup face of the same person. In addition, special attention is given to three local facial regions (foundation, eyes, and lips), each with its own pixel-level histogram loss function. Most GAN-based makeup transfer methods later adopted this region-based attention. An example of a BeautyGAN makeup transfer can be seen in Fig. 1.6.

These first attempts to augment a face with makeup were not particularly controllable. One could choose a makeup reference image, but it is not possible to adjust the intensity of the makeup application. BeautyGlow (Chen et al. 2019) attempts to detangle the latent space so that the source non-make image and the reference image are translated into distinct latent features, one corresponding to the face of the source

[6] The headshot photo is the courtesy of Matjaž Kljun. Used with permission. The images of the aged face were generated using the source code available at https://github.com/dawei6875797/Face-Aging-with-Identity-Preserved-Conditional-Generative-Adversarial-Networks.

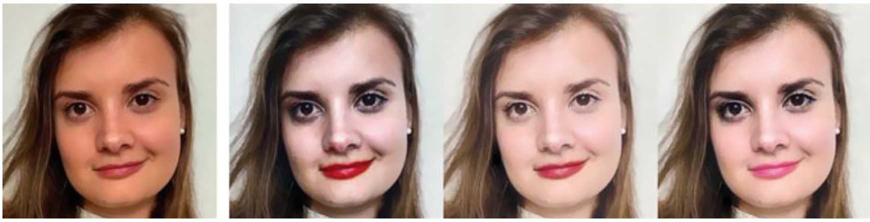


Fig. 1.6 Makeup applied[7] to a face with BeautyGAN. The left image is the real-world input image. The rest show a few different styles of virtual makeup application

image and the other to the makeup of the reference image. The two are then added together and translated into the output image. This approach gives some control over the intensity of makeup application. Moreover, such disentanglement allows better results with the reverse process of makeup removal.

Later methods attempt to optimize specific aspects of makeup transfer or removal. LADN (Local Adversarial Disentangling Network) (Gu et al. 2019) aims to overcome the limitation of focusing on local facial regions, such as the eyes and the mouth, and transfer or remove extreme makeup anywhere on the face. PSGAN (Pose and Expression Robust Spatial-Aware GAN) (Jiang et al. 2020; Liu et al. 2022) achieves better results when makeup is transferred between faces with different poses and expressions. It also allows partial makeup transfer, e.g., eye makeup from one and lipstick from another reference image. RAMT-GAN (Realistic and Accurate Makeup Transfer GAN) (Yuan and Zhang 2022) attempts to improve identity preservation and retain non-makeup regions, e.g., ears, hair, neck, etc., which earlier methods may modify in some cases. AM-NET (Aging-Net and Makeup-Net) implements makeup transfer for different age groups (Fang et al. 2022).

1.6 Generating 3D Models

In Augmented Reality, 3D representations are essential in achieving realism and deeper immersion. Objects that are inserted into a 3D world must be modelled. The end-user is often restrained to a set of predefined objects that can be used within a VR or an AR application. Two approaches were traditionally taken to create new 3D objects—model the objects manually or capture their shape by 3D scanning devices or depth sensors. Only in recent years, with the advances in deep learning and generative methods, have some researchers started to consider the possibility of generating 3D representations with GANs. Although we are probably still quite far

[7] The headshot photo is the courtesy of Marina Paldauf. Used with permission. The output images were generated using the source code available at https://github.com/Honlan/BeautyGAN.

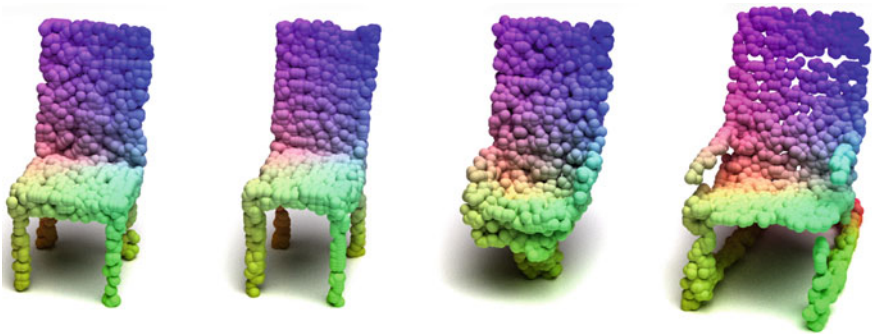


Fig. 1.7 Different types of chairs generated by the same 3D-GAN trained model[8]

from an advantageous solution, the general idea is intriguing. Let a GAN be trained on existing sets of 3D shapes and then generate new shapes of the same domain which have not yet been seen. This challenge has thus far been approached through the concept of *point clouds*.

A point cloud is a set of points in a 3D Euclidean space that represent surface-based geometry. It is a standard acquisition format used by 3D scanners and RGB-D sensors. The first known GAN architecture to randomly generate 3D objects was 3D-GAN (Wu et al. 2016). This architecture takes a 200-dimensional latent vector and maps it into a 64 $\times$ 64 $\times$ 64 voxel space. As such, it does not actually classify as a point cloud generator but still produces somewhat comparable results to actual point cloud GANs. The authors trained the model with different shape categories, such as chairs, sofas, and tables. The trained models can generate new shapes that do not exist within the training set. An example of 3D-GAN generated chairs can be seen in Fig. 1.7.

Instead of mapping into a voxel grid, latent-GAN (Achlioptas et al. 2018) maps a value from the latent space to an actual point cloud with a fixed number of 2048 points and is therefore considered by the community as the first GAN-based point cloud generator. This GAN architecture is coupled with a pre-trained autoencoder that encodes the point cloud before it is used with the GAN. This approach significantly simplifies the architecture of latent-GAN while still achieving realistically looking results. This architecture also proved useful for shape completions. For example, suppose a point cloud was obtained by scanning a partially occluded object; hence, some points are missing. The shape can then be completed through a pre-trained GAN model.

Only very recently, GAN-based point cloud generation started gaining attention in the context of AR. As pointed out by Lim et al. (2022), reduction in computational load is of crucial importance to make these types of deep learning architectures useful for AR content generation. Similar to latent-GAN, their GAN architecture uses

[8] The shapes were generated randomly using the official 3D-GAN source code available at https://github.com/zck119/3dgan-release.

Fig. 1.8 An example of applying a point cloud to an AR application[9]—choosing a piece of furniture to match the surroundings

a separate autoencoder to optimize the latent space, but they use a smaller encoding network based on the LDNCNN architecture (Zhang et al. 2021). This encoder can learn local geometric structures and avoid generating undesired artifacts. They envision using point cloud generators in VR and AR applications, where the demand for 3D models is high so that new shapes can be generated by smooth transition through the latent space. Figure 1.8 shows an example of applying a generated cloud point to an AR application as proposed by Lim et al. (2022).

A possible future trend may lie in generating meshes instead of point clouds since they are better suited for representing surfaces. Lin Gao and his colleagues demonstrated a GAN-based transfer of a 3D mesh deformation from one type of human body to another (Gao et al. 2018). Their later work (Gao et al. 2019) demonstrated the creation of new 3D mesh shapes through interpolation in the latent space. This latter approach did not employ a GAN architecture but variational autoencoders (VAs). Their work was not directly inspired by AR research trends, but a clear potential for 3D AR content generation can be seen.

1.7 Other GAN Use Cases in AR

Sometimes, researchers use GANs to tackle a specific problem, which is not a part of some research trends within AR. These are usually particular use cases—applications for some specific hardware, for conducting user studies, or simply for entertainment. This section briefly reviews some of the recent GAN applications in AR and VR that cannot be classified into one of the research areas indicated in this chapter.

[9] The mage was published in IEEE Transactions on Consumer Electronics under the Creative Commons License CC BY 4.0.

Several AR applications have been developed for entertainment purposes. A GAN-based 'deep fake' algorithm was used to entertain museum visitors in Georgia, USA (Wynn et al. 2021). The authors developed an AR application that provides an immersive user experience by animating the portraits at the museum's exhibition. The portraits could be made to appear on the screen as if they are speaking while the user listens to the audio through the headphones. A weather-predicting AR application used GAN architecture that was trained on radar observations spanning several months (Freeman 2020). The application was able to augment the camera stream with accurate rainfall pattern predictions up to an hour ahead. For a more immersive AR and VR user experience, head-mounted displays are often used. A study has been conducted on the feasibility of applying SRGAN to the Holo360 device (Feng et al. 2020). These types of devices currently support resolutions up to 4 K, while the resolution of the human eye is approximately 16 K, hence the motivation to upscale the display image while retaining its photo-realism.

Virtual reality is sometimes used in environments about which little information is available. Such was the case in training the localization system of a lunar rover, whose goal was to determine the rover's position on the map while it travelled inside a VR-simulated lunar terrain (Franchi and Ntagiou 2021). First, a GAN was trained to generate convincing images of the lunar environment, for which only 100 training images were used. Then, once the GAN had been trained, it worked similarly to a continuously expanding VR world, where the data from the VR environment was augmented with GAN-generated images.

An impressive case of a GAN-based style transfer in AR is the application called DeepTaste (Nakano et al. 2019). This application was developed to conduct a user study on vision-induced gustatory manipulation. The premise is that visual stimuli largely condition the sense of taste; therefore, by changing the appearance of the food through AR, the perceived taste of food can, to some degree, be manipulated. This finding could help people with dietary restrictions enjoy a wider range of tastes. The authors used style transfer on the live feed from a head-mounted device, which the user wore while eating. The color and the texture of the observed food were altered in real-time through a pre-trained StarGAN model. Figure 1.9 shows first-person experiences of AR gustatory manipulation. In the left image, the original food is occluded by a 3D model of a different type of food, while in the right image, StarGAN is used to change the appearance of the food. The research showed that the StarGAN method was preferred by the users.

Interaction within an AR scene primarily focuses on virtual objects inserted into the real world. When an interaction is attempted with real objects, 3D reconstruction is usually required to estimate the necessary structural information. Such reconstruction can be challenging and may require additional hardware for depth estimation. Generative methods can help reconstruct occluded parts of the scene or individual objects without the need for 3D reconstruction. In an AR experiment conducted by Son and Chang (2019), a GAN was trained on a database of 3D chair models, which were presented under different viewing angles. The trained GAN could then accurately approximate an image of a given chair as viewed under a given angle. The parts occluded in the original image were reconstructed in the rotated output by the

Fig. 1.9 Two examples of gustatory manipulation through AR.[10] Left: occluding the original food (somen noodles) with a 3D model (bowl of soup). Right: changing the appearance of the noodles through StarGAN

trained GAN. Odajima and Komuro (2021) attempted to reconstruct the whole scene, which was outside the camera's view. Their motivation was to construct a reflection of real-world surroundings on an inserted virtual object. The surroundings outside the field of view were extrapolated with the pix2pixHD GAN architecture.

1.8 Conclusion

There is a certain potential in the GAN deep learning architectures that the AR community can exploit and adapt to their particular needs. Although GANs have been around for almost a decade, they seem to have only recently made some impact on AR research, with many directions still open for exploration. A few isolated experiments and applications appeared in the last few years that explicitly connect the two research fields and offer specific AR use cases, while most other work discussed in this chapter deals with challenges that could potentially be used with AR applications and devices. Therefore, this chapter aims to provide an overview of the research area with highlighted ideas and potentials for AR, as well as to point out the impressive work done by those who developed GAN-driven AR applications.

Because of the popularity of GANs, many variations were proposed in a relatively short period. To a person not well acquainted with this development, the many different types of GANs encountered in different AR-related papers might be somewhat confusing. Furthermore, it often depends on when a particular GAN-driven AR application was developed. Only months later, an improved version of the used GAN architecture could be proposed, impacting further work on the application. For this reason, a detailed timeline of GAN evolution was presented and discussed, focusing on GAN methods that are fitted to generate images and videos.

[10] The images were contributed by the authors of DeepTaste. Used with permission.

We identified three research areas where GANs directly or indirectly impact the development of AR methods, namely: image composition, AR face filters, and generating 3D models. Image composition is an operation of combining visual elements from different sources into a photo-realistic composite image and is a such closely linked to AR. Existing GAN approaches focus predominately on object placement and shadow generation (Niu et al. 2021). Utilization of these methods in AR applications is straightforward—the video stream from the camera is augmented with virtual elements, which are harmonized with the background through GAN. Although the proposed algorithms achieve impressive results, not many AR applications can yet be found that utilize these approaches for some specific task rather than just as a proof of concept.

Face filters do not employ image composition techniques but aim to translate a real-world image to a virtual image while retaining photo realism. This concept adheres to the notion of image-to-image translation or style transfer, with face filters being a specific case. However, the current body of research in GAN style transfer for AR applications is almost exclusively focused on face filters, with three prevalent trends—changing facial features, face aging, and applying virtual makeup. Only a few exceptions can be found that employ GAN style transfer to other domains, e.g., the DeepTaste application (Nakano et al. 2019), where the appearance of food is changed instead of a face.

GANs have also been considered for generating 3D objects. We may expect this work to impact the development of immersive 3D applications eventually. This research area is still in its infancy, with only a handful of publications and seemingly no AR applications beyond mere ideas and proposals (Lim et al. 2022). It, nevertheless, has a certain potential and is worthy of keeping track of its further development. Thus far, GANs have been used to generate point clouds of 3D objects as an alternative to 3D scanning or sensing. Some successful attempts have also been made to synthesize meshes, which are better suited to represent surfaces than point clouds (Gao et al. 2019). Suppose that, in time, GANs become as proficient in generating realistic 3D content as they are currently in generating human faces (e.g., StyleGAN). AR applications could then operate with virtually unlimited supplies of 3D models without having to rely on either internal or cloud storage.

References

Achlioptas P, Diamanti O, Mitliagkas I, Guibas L (2018) Learning representations and generative models for 3D point clouds. In: Dy J, Krause A (eds) Proceedings of the 35th international conference on machine learning. PMLR, pp 40–49

Aldausari N, Sowmya A, Marcus N, Mohammadi G (2022) Video generative adversarial networks: a review. ACM Comput Surv 55. https://doi.org/10.1145/3487891

Antipov G, Baccouche M, Dugelay J-L (2017) Face aging with conditional generative adversarial networks. In: 2017 IEEE international conference on image processing (ICIP). IEEE Press, pp 2089–2093

Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: Precup D, Teh YW (eds) Proceedings of the 34th international conference on machine learning. PMLR, pp 214–223

Azadi S, Pathak D, Ebrahimi S, Darrell T (2020) Compositional GAN: learning image-conditional binary composition. Int J Comput Vis 128:2570–2585. https://doi.org/10.1007/s11263-020-01336-9

Bahmani S, Park JJ, Paschalidou D et al (2022) 3D-aware video generation

Brock A, Donahue J, Simonyan K (2019) Large scale GAN training for high fidelity natural image synthesis. In: International conference on learning representations

Brooks T, Hellsten J, Aittala M et al (2022) Generating long videos of dynamic scenes

Chen X, Duan Y, Houthooft R et al (2016) InfoGAN: interpretable representation learning by information maximizing generative adversarial nets. In: Lee D, Sugiyama M, Luxburg U et al (eds) Advances in neural information processing systems. Curran Associates, Inc.

Chen H-J, Hui K-M, Wang S-Y et al (2019) BeautyGlow: on-demand makeup transfer framework with reversible generative network. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 10034–10042

Choi Y, Choi M, Kim M et al (2018) StarGAN: unified generative adversarial networks for multi-domain image-to-image translation. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 8789–8797

Choi Y, Uh Y, Yoo J, Ha J-W (2020) StarGAN v2: diverse image synthesis for multiple domains. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 8185–8194

Clark A, Donahue J, Simonyan K (2019) Adversarial video generation on complex datasets. arXiv: Computer Vision and Pattern Recognition

Ekman P, Friesen WV (1978) Facial action coding system: a technique for the measurement of facial movement. Consulting Psychologists Press, Palo Alto, CA

Fang S, Duan M, Li K, Li K (2022) Facial makeup transfer with GAN for different aging faces. J Vis Commun Image Represent 85:103464. https://doi.org/10.1016/j.jvcir.2022.103464

Feng C-H, Hung Y-H, Yang C-K et al (2020) Applying Holo360 video and image super-resolution generative adversarial networks to virtual reality immersion. In: Human-computer interaction. Design and user experience: thematic area, HCI 2020, held as part of the 22nd international conference, HCII 2020, Copenhagen, Denmark, 19–24 July 2020, proceedings, Part I. Springer, Berlin, Heidelberg, pp 569–584

Franchi V, Ntagiou E (2021) Augmentation of a virtual reality environment using generative adversarial networks. In: 2021 IEEE international conference on artificial intelligence and virtual reality (AIVR), pp 219–223

Freeman J (2020) Content enhancement with augmented reality and machine learning. J South Hemisphere Earth Syst Sci 70:143–150

Gao L, Yang J, Qiao Y-L et al (2018) Automatic unpaired shape deformation transfer. ACM Trans Graph 37:1–15

Gao L, Yang J, Wu T et al (2019) SDM-NET: deep generative network for structured deformable mesh. ACM Trans Graph 38. https://doi.org/10.1145/3355089.3356488

Goodfellow I, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C et al (eds) Advances in neural information processing systems. Curran Associates, Inc.

Gu Q, Wang G, Chiu MT et al (2019) LADN: local adversarial disentangling network for facial makeup and de-makeup. In: 2019 IEEE/CVF international conference on computer vision (ICCV), pp 10480–10489

Gulrajani I, Ahmed F, Arjovsky M et al (2017) Improved training of Wasserstein GANs. In: Guyon I, Luxburg U von, Bengio S et al (eds) Advances in neural information processing systems. Curran Associates, Inc.

Hong Y, Niu L, Zhang J (2022) Shadow generation for composite image in real-world scenes. In: Proceedings of the AAAI conference on artificial intelligence, vol 36, pp 914–922. https://doi.org/10.1609/aaai.v36i1.19974

Hu X, Jiang Y, Fu C-W, Heng P-A (2019) Mask-ShadowGAN: learning to remove shadows from unpaired data. In: 2019 IEEE/cvf international conference on computer vision (ICCV), pp 2472–2481

Huang Z, Chen S, Zhang J, Shan H (2021) PFA-GAN: progressive face aging with generative adversarial network. IEEE Trans Inf Forens Secur 16:2031–2045. https://doi.org/10.1109/TIFS.2020.3047753

Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Bach F, Blei D (eds) Proceedings of the 32nd international conference on machine learning. PMLR, Lille, France, pp 448–456

Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 5967–5976

Jaderberg M, Simonyan K, Zisserman A, kavukcuoglu koray (2015) Spatial transformer networks. In: Cortes C, Lawrence N, Lee D et al (eds) Advances in neural information processing systems. Curran Associates, Inc.

Jiang W, Liu S, Gao C et al (2020) PSGAN: pose and expression robust spatial-aware GAN for customizable makeup transfer. In: IEEE/CVF conference on computer vision and pattern recognition (CVPR)

Kaneko T, Hiramatsu K, Kashino K (2017) Generative attribute controller with conditional filtered generative adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 7006–7015

Karras T, Aila T, Laine S, Lehtinen J (2018) Progressive growing of GANs for improved quality, stability, and variation. In: International conference on learning representations

Karras T, Aittala M, Laine S et al (2021) Alias-free generative adversarial networks. In: Ranzato M, Beygelzimer A, Dauphin Y et al (eds) Advances in neural information processing systems. Curran Associates, Inc., pp 852–863

Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 4396–4405

Karras T, Laine S, Aittala M et al (2020) Analyzing and improving the image quality of StyleGAN. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 8107–8116

Kikuchi K, Yamaguchi K, Simo-Serra E, Kobayashi T (2019) Regularized adversarial training for single-shot virtual try-on. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV) workshops

Kingma DP, Welling M (2014) Auto-encoding variational Bayes. In: Bengio Y, LeCun Y (eds) 2nd international conference on learning representations, ICLR 2014, Banff, AB, Canada, 14–16 April 2014, Conference Track Proceedings

Kossaifi J, Tran L, Panagakis Y, Pantic M (2017) GAGAN: geometry-aware generative adversarial networks. In: Proceedings/CVPR, IEEE Computer Society conference on computer vision and pattern recognition

Ledig C, Theis L, Huszár F et al (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 105–114

Lee D, Liu S, Gu J et al (2018) Context-aware synthesis and placement of object instances. In: Proceedings of the 32nd international conference on neural information processing systems. Curran Associates Inc., Red Hook, NY, USA, pp 10414–10424

Li T, Qian R, Dong C et al (2018) BeautyGAN: instance-level facial makeup transfer with deep generative adversarial network. In: Proceedings of the 26th ACM international conference on multimedia. Association for Computing Machinery, New York, NY, USA, pp 645–653

Lim S, Shin M, Paik J (2022) Point cloud generation using deep adversarial local features for augmented and mixed reality contents. IEEE Trans Consum Electron 68:69–76. https://doi.org/10.1109/TCE.2022.3141093

Lin C-H, Yumer E, Wang O et al (2018) ST-GAN: spatial transformer generative adversarial networks for image compositing. In: IEEE conference on computer vision and pattern recognition (CVPR)
Liu D, Long C, Zhang H et al (2020) ARShadowGAN: shadow generative adversarial network for augmented reality in single light scenes. In: The IEEE conference on computer vision and pattern recognition (CVPR)
Liu S, Jiang W, Gao C et al (2022) PSGAN++: robust detail-preserving makeup transfer and removal. IEEE Trans Pattern Anal Mach Intell 44:8538–8551. https://doi.org/10.1109/TPAMI.2021.3083484
Liu Y, Li Q, Sun Z (2019) Attribute-aware face aging with wavelet-based generative adversarial networks. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 11869–11878
Maas AL, Hannun AY, Ng AY (2013) Rectifier nonlinearities improve neural network acoustic models. In: ICML workshop on deep learning for audio, speech and language processing
Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv:1411.1784
Nakano K, Horita D, Sakata N et al (2019) DeepTaste: augmented reality gustatory manipulation with GAN-based real-time food-to-food translation. In: 2019 IEEE international symposium on mixed and augmented reality (ISMAR), pp 212–223
Niu L, Cong W, Liu L et al (2021) Making images real again: a comprehensive survey on deep image composition
Odajima S, Komuro T (2021) Reproduction of environment reflection using extrapolation of front camera images in mobile AR. In: 2021 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct), pp 300–303
Odena A (2016) Semi-supervised learning with generative adversarial networks
Odena A, Olah C, Shlens J (2017) Conditional image synthesis with auxiliary classifier GANs. In: Proceedings of the 34th international conference on machine learning, vol 70. JMLR.org, pp 2642–2651
Ohnishi K, Yamamoto S, Ushiku Y, Harada T (2018) Hierarchical video generation from orthogonal information: optical flow and texture. In: Proceedings of the AAAI conference on artificial intelligence, vol 32. https://doi.org/10.1609/aaai.v32i1.11881
O'Shea K, Nash R (2015) An introduction to convolutional neural networks
Pereira LT, Júnior WCR, Silva RLS (2021) Photorealism in mixed reality: a systematic literature review. Int J Virtual Real 21:15–29
Pumarola A, Agudo A, Martinez AM et al (2020) GANimation: one-shot anatomically consistent facial animation. Int J Comput vis 128:698–713. https://doi.org/10.1007/s11263-019-01210-3
Qiu H, Jiang Y, Zhou H et al (2022) StyleFaceV: face video generation via decomposing and recomposing pretrained StyleGAN3
Radford A, Metz L, Chintala S (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. In: 4th international conference on learning representations, ICLR 2016, San Juan, Puerto Rico, 2–4 May 2016, Conference Track Proceedings
Ratliff LJ, Burden SA, Sastry SS (2013) Characterization and computation of local Nash equilibria in continuous games. In: 2013 51st annual allerton conference on communication, control, and computing (Allerton), pp 917–924
Saito M, Matsumoto E, Saito S (2017) Temporal generative adversarial nets with singular value clipping. In: 2017 IEEE international conference on computer vision (ICCV), pp 2849–2858
Skorokhodov I, Tulyakov S, Elhoseiny M (2022) StyleGAN-V: a continuous video generator with the price, image quality and perks of StyleGAN2. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 3626–3636
Son M, Chang HS (2019) InteractionGAN: image-level interaction using generative adversarial networks. In: 2019 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct), pp 205–210

Sun Y, Tang J, Shu X et al (2020) Facial age synthesis with label distribution-guided generative adversarial network. IEEE Trans Inf Forens Secur 15:2679–2691. https://doi.org/10.1109/TIFS.2020.2975921

Tian Y, Ren J, Chai M et al (2021) A good image generator is what you need for high-resolution video synthesis. In: International conference on learning representations

Tripathy S, Kannala J, Rahtu E (2020) ICface: interpretable and controllable face reenactment using GANs. In: 2020 IEEE winter conference on applications of computer vision (WACV), pp 3374–3383

Tripathy S, Kannala J, Rahtu E (2021) FACEGAN: facial attribute controllable reenactment GAN. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision (WACV), pp 1329–1338

Tripathy S, Kannala J, Rahtu E (2022) Single source one shot reenactment using weighted motion from paired feature points. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision (WACV), pp 2715–2724

Tulyakov S, Liu M-Y, Yang X, Kautz J (2018) MoCoGAN: decomposing motion and content for video generation. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 1526–1535

Vondrick C, Pirsiavash H, Torralba A (2016) Generating videos with scene dynamics. In: Proceedings of the 30th international conference on neural information processing systems. Curran Associates Inc., Red Hook, NY, USA, pp 613–621

Wang T-C, Liu M-Y, Zhu J-Y et al (2018a) High-resolution image synthesis and semantic manipulation with conditional GANs. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 8798–8807

Wang Z, Tang X, Luo W, Gao S (2018b) Face aging with identity-preserved conditional generative adversarial networks. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 7939–7947

Wang Y, Bilinski P, Bremond F, Dantcheva A (2020) G3AN: disentangling appearance and motion for video generation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)

Wu J, Zhang C, Xue T et al (2016) Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. In: Advances in neural information processing systems, pp 82–90

Wynn N, Johnsen K, Gonzalez N (2021) Deepfake portraits in augmented reality for museum exhibits. In: 2021 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct), pp 513–514

Yang H, Huang D, Wang Y, Jain AK (2018) Learning face age progression: a pyramid architecture of GANs. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 31–39

Yuan Q-L, Zhang H-L (2022) RAMT-GAN: realistic and accurate makeup transfer with generative adversarial network. Image Vis Comput 120:104400. https://doi.org/10.1016/j.imavis.2022.104400

Zhan F, Zhu H, Lu S (2019) Spatial fusion GAN for image synthesis. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 3648–3657

Zhan F, Lu S, Zhang C et al (2020) Adversarial image composition with auxiliary illumination. In: Proceedings of the Asian conference on computer vision (ACCV)

Zhang H, Goodfellow I, Metaxas D, Odena A (2019) Self-attention generative adversarial networks. In: Chaudhuri K, Salakhutdinov R (eds) Proceedings of the 36th international conference on machine learning. PMLR, pp 7354–7363

Zhang L, Wen T, Min J et al (2020) Learning object placement by inpainting for compositional data augmentation. In: Vedaldi A, Bischof H, Brox T, Frahm J-M (eds) Computer vision—ECCV 2020. Springer, Cham, pp 566–581

Zhang K, Hao M, Wang J et al (2021) Linked dynamic graph CNN: learning through point cloud by linking hierarchical features. In: 2021 27th international conference on mechatronics and machine vision in practice (M2VIP), pp 7–12

Zhou H, Ma R, Zhang L-X et al (2021) SAC-GAN: structure-aware image composition

Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks

Chapter 2
Augmented Reality User's Experience: AI-Based Data Collection, Processing and Analysis

Yevgeniya Sulema, Andreas Pester, Bertrand Laforge, and Frederic Andres

Abstract This chapter presents an AI-based approach to the systemic collection of user experience data for further analysis. This is an important task because user feedback is essential in many use cases, such as serious games, tourist and museum applications, food recognition applications, and other software based on Augmented Reality (AR). For AR game-based learning environments user feedback can be provided as Multimodal Learning Analytics (MMLA) which has been emerging in the past years as it exploits the fusion of sensors and data mining techniques. A wide range of sensors have been used by MMLA experiments, ranging from those collecting students' motoric (relating to muscular movement) and physiological (heart, brain, skin, etc.) behaviour, to those capturing social (proximity), situational, and environmental (location, noise) contexts in which learners are placed. Recent research achievements in this area have resulted in several techniques for gathering user experience data, including eye-movement tracking, mood tracking, facial expression recognition, etc. As a result of user's activity monitoring during AR-based software use, it is possible to obtain temporal multimodal data that requires rectifying, fusion, and analysis. These procedures can be based on Artificial Intelligence, Fuzzy logic, algebraic systems of aggregates, and other approaches. This

Y. Sulema (✉)
Department of Computer Systems Software, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
e-mail: sulema@pzks.fpm.kpi.ua

A. Pester
Faculty of Informatics and Computer Science, The British University in Egypt, Cairo, Egypt
e-mail: andreas.pester@bue.edu.eg

B. Laforge
Laboratoire de Physique Nucléaire et des Hautes Energies (LPNHE), Sorbonne Université, Paris, France
e-mail: laforge@lpnhe.in2p3.fr

F. Andres
Digital Content and Media Sciences Research Division, National Institute of Informatics, Tokyo, Japan
e-mail: andres@nii.ac.jp

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_2

chapter covers theoretical and practical aspects of handling AR user's experience data, in particular, MMLA data. The chapter gives an overview of sensors, tools, and techniques for MMLA data gathering as well as presenting several approaches and methods for user experience data processing and analysis.

2.1 Introduction

Augmented Reality (AR) technology offers a new dimension to serious games, touristic and museum applications, food recognition applications, and other similar software. At the same time, ensuring the high quality of AR-based software requires the monitoring of the user's experience. For AR game-based learning environments user feedback can be provided as Multimodal Learning Analytics (MMLA) which has been emerging in the past years as it exploited the fusion of sensors and data mining techniques.

To collect MMLA data, we can use several devices, tools, and techniques which will complement the raw educational data coming from a digital educational application. For instance, one can use a camera whose image stream is analysed online by an AI algorithm to provide information about the body or face language of a learner. One can also use distinct types of bio-oriented sensors measuring blood-pressure, heart pulse, breath rhythm and importantly neuro-devices reporting about different signals inside the brain (some of them being only accessible on specific research platforms such as NeuroSpin (2022) but some devices can now also be accessible in a class). An important set of useful sensors are also available in the hands of the learners, namely their mobile phone and modern digital watches. Such connected devices are of great interest because they provide data collection options at large scale and low price.

Data presenting MMLA requires processing that can be based on Artificial Intelligence (AI) methods. AI can be used for real-time feedback to the learner, for aggregation of information about the learner's behaviour in-situ, provide summarised knowledge after the learning session to guide the learner in different learning situations, to update the digital learner model, and to give feedback to the developers of the AR systems. The AI methods can predict and classify learning situations as learner behaviour as well.

Thus, in this chapter, we present important aspects of AR-software user experience data collecting, processing, and analysis based on AI. The rest of the chapter is organised as follows. In Sect. 2.2, we analyse recent related work. In Sect. 2.3, we consider use cases and scenarios for serious AR-based games. In Sect. 2.4, we focus on systemic data collection. In Sect. 2.5, we present a range of capturing devices and tools that enable the monitoring of user experience and activity parameters. In Sect. 2.6, we discuss AI-enhanced data processing methods. In Sect. 2.7, we consider the creation of a complex model of an AR-software user. Finally, in Sect. 2.8, we summarise the presented approach.

2.2 Related Work

Recent research demonstrates special attention to the implementation of innovative approaches such as AI and digital twins in the educational area.

Cowling and Birt (2020) give an overview of the most common approaches for Mixed reality multimodal learning analytics.

Ochoa and Worsley (2016) analyse how and why audio and video can be used in non-technology-centred learning environments to collect signals (data) about the learners' actions and interactions. Several variants of multimodal learning analytics are discussed.

Di Mitri et al. (2018) conducted a literature survey of experiments using multimodal data for multimodal learning analytics. They analysed input multimodal data and the related learning theories. Finally, they formulated a Multimodal Learning Analytics Model with three objectives:

- Enhancing the feedback in a learning context.
- Combining several machine learning algorithms with multimodal data.
- Aligning the used machine learning and learning science terminology.

Marcel (2019) provides a field study about different mobile Augmented Reality applications and learning objects for higher education on the mobile AR platform HP Reveal.

Spikol et al. (2018) provide an experimental study on how to find the most notable features for project-based learning set-ups using supervised machine learning for multimodal learning analytics.

Guo et al. (2022) give an overview of actual trends in the use of AI to enhance the metaverse project. These trends are not only related to Metaverse but have an impact on the area of computer vision and natural language processing, as the sensor architecture as well.

Tao and Xu (2022) analyse the application of digital twin technology in different areas including aviation and education. The authors come to the conclusion that digital twin technology has potential for the field of education. To prove this, they present an example of constructing a digital twin for education based on HTC Vive.

Berisha-Gawlowski et al. (2021) considers the potential of the digital twin from two perspectives: (1) application of digital twin technology to control and manage a cyber-physical system, and (2) representation of humans in the virtual world. Considering the application of digital twin technology in the field of education, the authors raise questions related to both technological and methodological aspects of using digital twins for educational purposes. To illustrate their vision, the authors refer to the example of team learning.

Rudra (2022) analyses the possibility of digital twin technology becoming the next step in immersive learning. The author provides ideas on enhancing learning in medicine and chemistry by employing digital twin technology in combination with AR.

The analysis of these and other recent papers allows us to draw the conclusion that these technologies have a high potential for AR-based educational game applications as well.

2.3 Use Cases for Educational AR Games

Given the current constraints on energy consumption, while targeting the largest possible usage of an educational game, it is important to design AR games with the minimum amount of technology that allows it to meet its educative objectives. A question is then to identify cases where AR is worth implementing in an educative game. On the other hand, educational games including AR are an effective way to collect data for MMLA.

The first obvious use case is when a game is using the camera of a mobile phone or a computer to enrich the scene with live information overlaid with the real image taken by the player. In terms of education, this can be a modality to identify in the scene some objects of interest to play with or to bring complementary information to the learner. Depending on the amount of such information that a learner may need, can for instance gain clues about the autonomy of the learner to achieve a specific task. Let us imagine for instance a game that teaches security measures inside a laboratory or a firm's workshop: "how many critical situations can be identified by the learner in full autonomy?" and "how much help should be brought to detect some missed cases?" are questions that are of high relevance to assessing the competence of the learner.

The second case of interest is provided by a game aiming at discovering the full history of a painting and how this history was learned. In such a game, the learner could try to inspect the real painting with his camera while having some extra tools to perform dedicated analysis (it could be a magnifying glass showing the x-ray picture of a part of the painting or a simulated spectrum that is acquired using a proton or neutron beam to inspect the composition of the painting, or it could be a tool that reveals the modifications brought by the conservators over time, etc.). Such complementary information is of importance in presenting art pieces in a very new way.

The third case of importance is given by the interesting options that are provided by AR in terms of discovering levels of reality. Our immediate environment is indeed not what it seems to be as seen by our human senses. In a class or a museum of sciences, it is remarkably interesting to have the ability to change one's scale of observation, for instance, to go to the molecular level by providing a simulation at a scale chosen by the learner. This can be particularly useful to teach physics or chemistry. Consider, for example, how difficult it is to discover what a magnet is in terms of atomic properties and how AR can support this learning task.

2.4 Systemic Data Collection

Systemic data collection in the context of digital education aims at collecting nominative data that will be used by learning analytics tools to provide immediate or delayed personalised feedback to learners while also providing large-scale data samples to support meaningful research and to train complex AI algorithms. Research can be devoted to demonstrating interesting features of specific pedagogical approaches such as AR-based game training but can also deal with the difficult problem of connecting the dots between a model of the learner, a curriculum, and the effective learning achieved as the result of this curriculum.

Large-scale data collection means going beyond a single educational institution scale and requires a specific design that can be generalised to many educational structures. The need for nominative data handling at the level of each institution is best achieved using a standard data format such as Learning Record Stores (LRS) using the xAPI ecosystem (Šimić et al. 2019; Rustici Software 2022). On the other hand, the usage of data from different institutions for research purposes is also needed and requires interoperable data formats and tools. This second need also requires installing a Data Processing Unit (DPU) that simultaneously has direct access to the local data and can communicate with other DPUs. With such a design, it is possible to analyse data locally but also request data processing at other sites. The privacy of data is of course questioned by such a large-scale analysis. This question can be solved by adding a software layer (API) that guarantees that processing outputs at any DPU can only be transmitted to other DPUs if they are aggregated results from the processing of a collection of individual information. An obvious example would be the output of a numerator and a denominator at each DPU level to reconstruct a percentage at the requesting DPU level. A more elaborated one could be the transmission of a mean gradient calculated over a set of learners at a given site transmitted to the requesting DPU to train a Machine Learning algorithm. Such a system can be viewed as an edge processing scheme, and it is then easy to expand to a larger scale. The CPU power will be used where the data is, and only aggregated results would be transmitted over the network. This guarantees the privacy of the data at each site and is potentially more efficient in terms of storage and network traffic than extracting anonymous data from each site to be centralised on a research cluster. It should be noticed that such a design is in line with current efforts made in Europe (Gaia-X Hub Germany 2022; Gaia-X European Association for Data and Cloud AISBL 2021; Prometheus 2022) to develop edge computing, interoperability of data, and sovereign cloud computing.

It is important to mention that the xAPI standard is unfortunately so flexible that further standardisation needs to be done to collect data to be used in an interoperable way. Going beyond the "actor + verb + complement" specification of xAPI is a necessity to allow global analyses of heterogeneous curricula in a coherent way while still allowing a detailed and specific analysis of a particular curriculum or educational tool. The xAPI ecosystem provides semantic versioning of statement vocabulary that allows Clients and LRSs to remain interoperable as the specification changes but there is much more to define in the context of many specific educative activities, especially for serious games and AR activities.

In the context of the Ikigai (2022) serious gaming platform (Fig. 2.1) developed in France by a consortium of higher education institutions and handled by a non-profit organisation known as Ikigai-Games for Citizens, a design for such specifications has been elaborated defining a hierarchy of xAPI statements that goes from global to specific concerns. Consequently, each game has to declare for instance a "start of game" and "end of game" that allows counting the number of games that have been played by learners regardless of the type of games. Another level of mandatory information requires the game to report specific activities that can be linked to a category of games, for instance, the registration of each shoot in a first-person shooter game. This specification also allows game-specific xAPI statements. On purpose, this hierarchy allows researchers to develop analyses with different granularities in a coherent way either aggregating data from different games to provide for instance global statistics about the overall usage of games in the educational activities of learners or choosing a lower granularity to provide statistical information of the impact or usage of a specific game type on student performance up to game specific information. This multiscale approach can be extended to other types of digital educative activities and requires a community effort to specify for instance the generic information that is needed to study AR systems such as an AR serious game. Learning analytics-based platforms should then have access to a hierarchy of AR system related xAPI statements when defining the type of data, they provide through AR activities. The same information should be also available to data analysts. Moreover, one needs a mechanism that guarantees that the collected information collected at two different periods can be analysed simultaneously. This feature will be optimally implemented through a versioning system of the AR xAPI information hierarchy available through a dedicated server used by all platforms. With such a system any learning analytics-based platform will be able to define the type of data that it provides in a way that will also be immediately available to data analysts through the same server. Interoperability of this metadata will be further simplified by a possible unique request to this centralised server providing the description of the available information in a dataset using a specific version of the hierarchy of xAPI statements.

In terms of infrastructures, such a large-scale system, designed to provide personalised guidance to students based on their individual data, also need to be thought of in association with a complementary pedagogical resource management system. This latter provides a mechanism to find content to recommend to the learners. Given the ongoing evolution towards a competence approach at the higher education level, a remarkably interesting framework is given by the Memorae platform developed at Université Technologique of Compiègne in France (Atrash et al. 2015). This platform implements the idea that a curriculum can be described as a tree or a graph of related competencies by a teacher. Each node is related to a competence while a graph connection models the relative dependence between two individual competencies. This map can then be made available to a learning community as an interactive tool allowing one to enter at the level of each node into a forum dedicated to its related

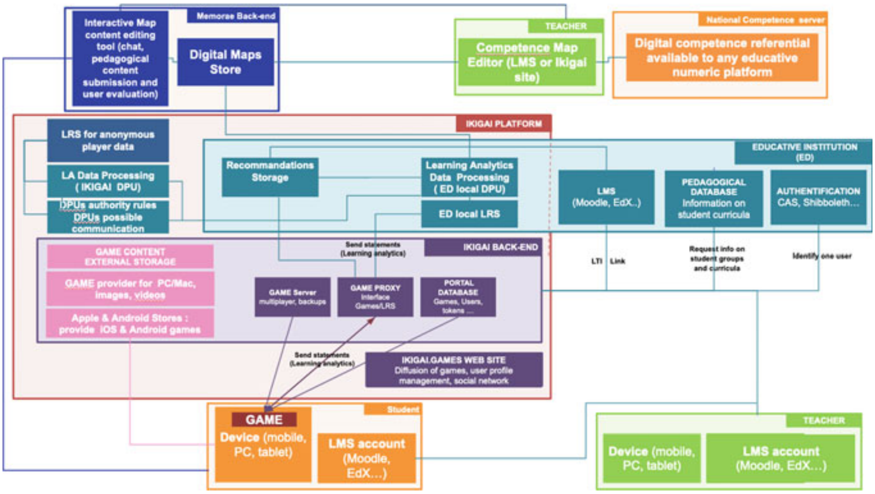


Fig. 2.1 Ikigai platform architecture

competencies. It further enables recommending content to work this competence and to evaluate the recommended contents (a PDF-file, a book, a film, a game, a simulation, etc., related to that competence). Beyond live or asynchronous exchanges between learners, this interactive map is then an educative content database that can be further used as a data structure to recommend content to be used by a student in the context of pedagogical advice provided by an AI-based recommendation system. The same map can be used to project single player data to identify its level of mastering of the various competencies defined in the map and then used to make those recommendations.

2.5 Monitoring Parameters Capturing Hardware and Tools

Multimodal analysis has demonstrated effectiveness in studying and modelling several human–human and human–computer interactions. In this section, we review the role of parameter capturing hardware and multimodal tools in the service of studying complex learning analytics environments.

One important set of parameters represents the affective factors (e.g., motivation, stress, or flow) and can be measured in diverse ways (e.g., clicks, postings, messages, views, writes and likes). Several studies (Kumar et al. 2022; Siddharth and Sejnowski 2022) demonstrated that there is a correlation between certain types of sensor data, so called biomarkers (e.g., heart rate), and higher-level states of persons that are relevant for learning (e.g., emotion, including anxiety). In the domain of IoT, a digital biomarker is defined as digitized data collected from learners via IoT devices (Nam et al. 2019). Learners' biomarkers are collected with three types of affordable sensors: an eye-tracker, an electroencephalogram, and a camera that monitors variations in different modalities, e.g., speaking, gesturing, gazing, typing (Lazar et al. 2017). The use of sensor technology enables the collection of learners' physiological and behavioural data. These data are the tracking of all the attitudes, interests, and motivations of the learners toward the educational experience, summarised as affective learning (Bamidis 2017).

To monitor the student's activity during a learning session, we can also use devices that capture and track the movement of the students' eyes: features including pupil size, saccade, fixations, velocity, blink, pupil position, fixation, electrooculogram, and gaze point (Eye Square 2022; Sangu 2020; Wang et al. 2018). Collected eye-tracking data serves to identify and analyse patterns of visual attention of individuals during learning experiences. There are four types of eye trackers:

- Tower mounted type (INITION London 2022; University of Edinburgh 2022)
- Screen-based type (Grossman et al. 2019; Biopac Systems Inc. 2022)
- Head mounted type (Sugano and Bulling 2015; Cognolato et al. 2018; Melnyk et al. 2022; Franchak and Chen 2022)
- Mobile type (Callemein et al. 2019; Müller et al. 2019; Liu et al. 2019).

Innovative approaches to include mood flow in learning experiences have been using non-intrusive mood-related sensors (Nashed et al. 2021). In general, Multimodal Data Processing uses several channels and spaces for monitoring. Di Mitri et al. (2018) distinguish an input space (different sensor signals) and a hypothesis space (latent attributes). The difference is in observability: input signals can be observed directly, and latent parameters can be concluded based on different inference models. The parameters monitored in these channels are mostly person and application centralised. It takes less into consideration the requirements of large-scale data collection described in Sect. 2.4.

One question is also, where the data are processed: in edge computing directly, on the mobile device, or centralised in the cloud. This needs standardisation. The standardisation of data processing should be in accordance with the standardisation, granulation, and semantic versioning of the systemic data collection.

For AR applications a smaller set of parameters can be used, including direct spatial data, motoric and behavioural data of the user of the learning game, and data from the interoperability of gamer AR user groups.

The implementation of the feature fusion strategy of deep neural networks for multimodal activity recognition faces the challenge of integrating features of different scales for better performance, as it has been shown in (Dai et al. 2021). The authors of (Münzner et al. 2017) classified the feature fusion strategy in four categories, initially applied in CNN-based architectures:

- Feature Fusion
- Sensor-based Fusion
- Axis-based Fusion
- Shared-filter Fusion.

Research enhancement has been focussing on the generalisation of feature fusion to any deep learning architectures (Schweizer et al. 2021; Siegfried and Odobez 2022; Sugano and Bulling 2015).

2.6 AI-Enhanced Data Processing Methods

For the Data Processing of the MMLA data, different methods can be used: Machine and deep learning classification methods, fuzzy logic predictions of linguistic variables, but also a special algebra of aggregates, developed for the handling of multimodal data in general.

The methods of processing data can be characterised by the algorithms used, the location of processing (edge or cloud-computing), the explainability of the results, and the used output.

From an algorithmic point of view, the methods differ in the question of whether they are data-driven (machine and deep learning) or rule-driven (fuzzy logic, algebra). In the first case the collected, cleaned, and fused data are used to train and test a machine learning model (for the user or learner or for the relation game input and learning outcome). Different methods had been used, for example, regression models, Bayes naive learning, support vector machines with linear and Gaussian kernels, and distinct types of artificial neural networks. In the last two years primarily transformer algorithms for language-based models have been used. The problem of transformers is that they need a huge amount of memory and computational resources and cannot be used directly by small or medium scale platforms or devices. In the future methods, which can take into consideration the multiple connections between the multimodal input data, like graph neural networks can be used. If multimodal data are considered as multivariate sequential data, methods used for time series are also an interesting approach.

To ensure consistency of data processing, the algebraic system of aggregates (ASA) can be used in combination with the fuzzy logic approach (Sulema and Kerre 2020). The ASA offers a way of aggregating data sets that can be used for the synchronisation and consolidation of data describing the same object of research (a learner in our case) if the data streams are received from multiple sources (sensors, devices, tools, etc.). It is especially important in the case of temporal multimodal data processing. An aggregate in the ASA is defined as a mathematical object that has the following distinguishing features:

- An aggregate consists of data tuples that are ordered; the order of tuples is important for the result of operations on aggregates (logical operations and ordering operations).
- A tuple in the aggregate can consist of values of the same type, values of several types, or tuples of values.

These features enable multiple variants of time-wise data processing in an effortless way.

Also, important is the location of data processing. Most of the AI applications are processed on a powerful computer with a GPU processor with enough RAM or directly on the cloud by different machine learning service providers. All Large Language Models are cloud applications. But AR applications are running on a smartphone or tablet and so far in a lot of cases the data should be processed directly on the device, otherwise, the latency time would be too long. Therefore, the algorithms should be dressed in such a way that they fit the device hardware. Some smartphones are already able to handle larger models, but AI on the edge is still an extensive research option. Sometimes a CNN accelerator is implemented directly in the visual recognition sensor unit, such as ShiDianNao, RedEye, and other applications (see in Guo et al. 2022). Another approach is to compress the models or to work with edge inference machines. A third way to solve the computing bottleneck is the development of in-memory computing, SRAM-based, DRAM-based, and novel memory technologies. Applying AI to more complex AR will require solving the problem of how to increase the capacity to run larger-scale Deep Learning models directly on a device.

A major challenge in Machine and Deep Learning and for AI applications to AR is the explainability of the results. On the one hand, the MMLA results are used to model the learners' behaviour, on the other side, with the help of modern generative methods like stable diffusion, it will be possible to create AR objects directly during the learner's communication with the learning object and the learning community. This can have an impact on the learner's behaviour, which we should be able to understand. The used deep learning methods are seen as a black box and the results so far require explanation. For classical machine learning, with the methods SHAP and LIME, explainability can be supported (Gashi et al. 2022), but for deep learning this is still a question under research.

The output of the data processing in MMLA can be represented as a classification of the learner/user/feedback type, detection of abnormal behaviour, but can also include sorting and compressing of the features of the input data. If Graph Neural Networks are used, a new node (virtual object) or edge (for the learner unbelievably valuable information), can be predicted, but also learning phases can be sorted by score.

With these methods, the mental and emotional situation of the learners can be learned and used for learning analytics and as input to the user's digital twin.

2.7 AR-Game User's Digital Twin

An educational AR-game user's digital twin is a continuously updating model that reflects individual features of the game player: a level of involvement in the study process, learning strategy and behaviour, etc. This model can be used for providing the learner with recommendations on the individual study trajectory for achieving better progress in learning. A collection of digital twins can be used for gathering generalised anonymized statistics about the AR-game quality, typical strategies, and learners' behaviour to help the game developers in game improvement.

The user's digital twin is represented through data obtained on the level of systemic data collection discussed in Sect. 2.4 of this chapter.

Beyond the data collection infrastructure itself, the hierarchy of information must be flexible enough to handle heterogeneous types of information. This can be achieved by defining a list of devices that provide the information, with a description of their capabilities as metadata, and add them as actors in the reporting of the actual scene of the games: "the thermometer reports that the face temperature of the learner is 38 °C", "the camera has been turned to infrared mode", etc. should be possible statements in the data collection. This type of information is complementary to the usual learner action or achievement information discussed previously and reported as a statement involving the learner as an actor. Extending the nature of available information to be processed by AI algorithms is a key point of improvement of current learner experience data collection since they should bring much more contextual information that allows a better understanding of the conditions under which an educative objective is met or not. All is not always a question of digesting formal information. Known facts about how a learner achieves memorising is very dependent on the psychological state of his/her brain, suggesting the use of neuroscience devices in the learner experience reporting.

To construct the user's digital twin based on data to be received from different devices and tools, we can use a formal specification to represent the important level of the user's model. This specification can be considered as a semantic model of the data describing the user as an object of research. Figure 2.2 shows an example of the semantic model for the user's digital twin creation.

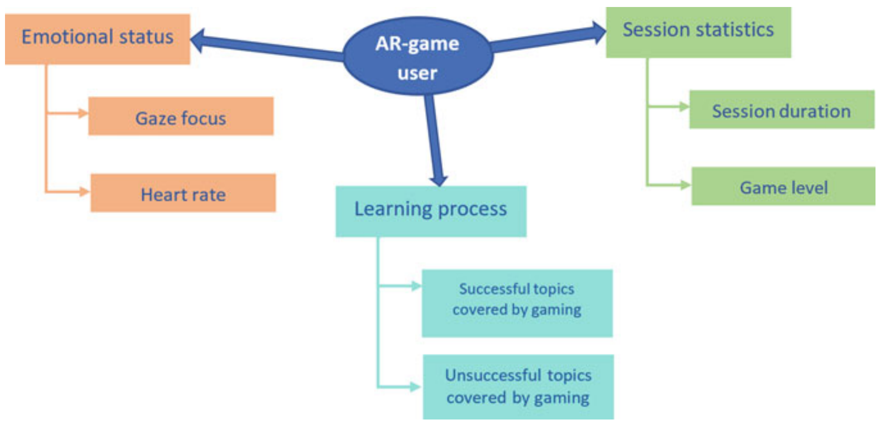


Fig. 2.2 AR-game user's semantic model

To develop the digital twin, we need to implement this model as a software tool that receives data streams according to the defined semantics, fuses them based on the synchronisation rules, and provides AI-based analysis and processing of aggregated data. The architecture of such software is shown in Fig. 2.3.

This architecture implements the following logic of using the AR-game user's digital twin. During an AR-gaming session, the user's parameters defined by the AR-game user's semantic model are being monitored using a set of devices and tools. These data streams come to the Data Fusion Module where they are synchronised and aggregated according to the data synchronisation rules. The data aggregate is stored in the cloud Data Storage. Using the AR-game user's individual data, the AI-based Analytics Module, through the Visualisation Module, provides the learner with analytics on study progress. It also provides the Individual Learning Recommendation System with the data to be used for helping the learner in forming an individual learning trajectory aimed at achieving higher quality of learning.

A teacher can use the system to get generalised information about learners. The queries for consolidated data analytics are processed by the Query Processing Module, which gets data from the cloud Data Storage and sends the requests for data processing to the AI-based Analytics Module. The resulting analytics is demonstrated by the Visualisation Module. These analytics help the teacher to improve the methodology of the AR-game. Besides, the AI-based Analytics Module provides feedback to the AR-game developer through the Game Development Feedback mechanism, enabling the advancement of the game development.

Thus, the general framework of AR-game user's digital twin software implementation and exploitation includes:

- The game development and advancement use case.
- The game efficiency analysis use case.
- The studying process enhancement use case.

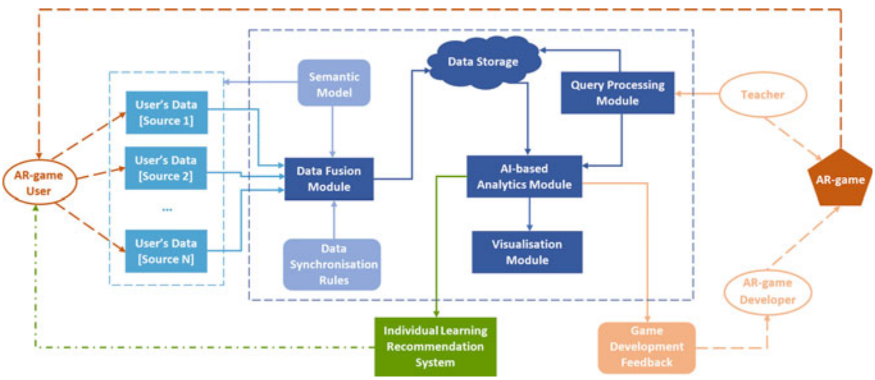


Fig. 2.3 AR-game user's digital twin software architecture in the general context

In the game development context, the AR-game user's digital twin can be used for obtaining analytics for a better understanding of the necessary advancement of the game: its logic, AR elements, controls, etc. In the game efficiency and learning strategies context, the digital twin can be used for the improvement of the teaching methodology. In the study context, the digital twin can be used for individual learning trajectory elaboration and recommending to the learner.

2.8 Conclusion

The systemic approach developed to collect data efficiently and at large scale in the domain of educative games reported here can be extended to AR games and other digital learning tools in a straightforward way and provides a minimal mechanism to handle educative recommendations for all the disciplines built by the learning communities themselves. By including heterogeneous types of data inside a global hierarchy of content, this system will allow the educative community to build complex models of learners and of their related curricula. Such a systemic approach will break the current limits that prevents extensive usage of AI in education. This system will further enhance the usage of data because it allows both a nominative usage of the data at the scale of a local institution and an anonymised access to these data for large scale research in education.

Multi-modal sensory data captures cognitive, motivational, and emotional behaviour of the learners during the learning processes. All those perceptions enable us to define the learners' affective learning.

AI enhances AR application for games and education in three aspects: (1) smart user feedback prediction and classification, (2) online learning (training) of the AI model based on the systemic data collection and (3) and predictions for similar user and use-cases. In the future, AI for AR research should concentrate on the AR object creation process based on modern technologies like text/speech to image, image to image diffusion, NLP for steering the scenes, and knowledge graphs for interoperational metadata exchange and processing.

The digital twin software is a useful tool that is aimed at the improvement of three important aspects: (1) AR-game design at the systemic level, (2) game exploitation teaching methodology, and (3) game-based studying. Future research in this context can be focused on the elaboration of the standardisation of AR-game user's digital twin software and accumulation of AR-based game exploitation best practices.

References

Atrash A, Abel MH, Moulin C (2015) Notes and annotations as information resources in a social networking platform. Comput Human Behav Elsevier 51:1261–1267

Bamidis PD (2017) Affective Learning: Principles, Technologies, Practice. In: Frasson C, Kostopoulos G (eds) Brain function assessment in learning. BFAL 2017 (Lecture notes in computer science), vol 10512. Springer, Cham. https://doi.org/10.1007/978-3-319-67615-9_1

Berisha-Gawlowski A, Caruso C, Harteis C (2021) The concept of a digital twin and its potential for learning organizations. In: Ifenthaler D, Hofhues S, Egloffstein M, Helbig C (eds) Digital transformation of learning organizations. Springer, Cham. https://doi.org/10.1007/978-3-030-55878-9_6

Biopac Systems Inc (2022) Robust eye tracking in a compact module. https://www.biopac.com/product/eye-tracking-eyetech/. Accessed 27 Oct 2022

Callemein T, Van Beeck K, Brône G, Goedemé T (2019) Automated analysis of eye-tracker-based human-human interaction studies. In: Kim K, Baek N (eds) ICISA 2018 (Lecture notes in electrical engineering), vol 514. Springer, Singapore. https://doi.org/10.1007/978-981-13-1056-0_50

Cognolato M, Atzori M, Müller H (2018) Head-mounted eye gaze tracking devices: an overview of modern devices and recent advances. J Rehabilit Assistive Technol Eng 5. https://doi.org/10.1177/2055668318773991

Cowling MA, Birt JR (2020) Mixed reality multimodal learning analytics. Encycl. Educ. Innov. 2020. https://doi.org/10.1007/978-981-13-2262-4

Dai Y, Gieseke F, Oehmcke S, Wu Y, Barnard K (2021) Attentional feature fusion. In: 2021 IEEE WACV2021, pp 3559–3568. https://doi.org/10.1109/WACV48630.2021.00360

Di Mitri D, Schneider J, Specht M, Drachsler H (2018) From signals to knowledge: a conceptual model for multimodal learning analytics. J Comput Assist Learn 34:338–349. https://doi.org/10.1111/jcal.12288

Eye Square (2022) Head mounted eye tracking https://www.eye-square.com/en/headmounted-eye-tracking/. Accessed 27 Oct 2022

Franchak JM, Chen Y (2022) Beyond screen time: Using head-mounted eye tracking to study natural behavior, Editor(s): Rick O. Gilmore, Jeffrey. J Lockman, Adv Child Develop Behav JAI 62:61–91. https://doi.org/10.1016/bs.acdb.2021.11.001

Gaia-X European Association for Data and Cloud AISBL (2021) Gaia-X architecture document. https://docs.gaia-x.eu/technical-committee/architecture-document/latest/. Accessed 27 Oct 2022

Gaia-X Hub Germany (2022) What is Gaia-X? https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html. Accessed 27 Oct 2022

Gashi M, Vuković M, Jekic N, Thalmann S, Holzinger A, Jean-Quartier C, Jeanquartier F (2022) State-of-the-art explainability methods with focus on visual analytics showcased by Glioma classification. BioMedInformatics 2:139–158. https://doi.org/10.3390/biomedinformatics2010009

Grossman RB, Zane E, Mertens J et al. (2019) Facetime versus screentime: Gaze patterns to live and video social stimuli in adolescents with ASD. Sci Rep 9:12643. https://doi.org/10.1038/s41598-019-49039-7

Guo Y, Yu T, Wu J et al. (2022) Artificial intelligence for metaverse: a framework. CAAI Artif Intell Res 1(1):54–67. https://doi.org/10.26599/AIR.2022.9150004

Ikigai (2022) Games for Citizens. https://www.gfc.ikigai.games/?lang=en. Accessed 27 Oct 2022

INITION London (2022) Motion capture & tracking sensomotoric instruments iView X Hi-speed. https://www.inition.co.uk/product/sensomotoric-instruments-iview-x-hi-speed/. Accessed 27 Oct 2022

Kumar A, Sharma K, Sharma A (2022) MEmoR: a multimodal emotion recognition using affective biomarkers for smart prediction of emotional health for people analytics in smart industries. Image vis Comput 123:104483. https://doi.org/10.1016/j.imavis.2022.104483

Lazar J, Feng JH, Hochheiser H (2017) Research methods in human-computer interaction. Cambridge, MA: Morgan Kaufmann. ISBN: 978-0-12-805390-4

Liu ZX, Liu Y, Gao X (2019) Using mobile eye tracking to evaluate the satisfaction with service office. In: Marcus A, Wang W (eds) Design, user experience, and usability. Practice and case studies. HCII 2019, (Lecture notes in computer science). Springer, Cham, p 11586. https://doi.org/10.1007/978-3-030-23535-2_14

Marcel F (2019) Mobile augmented reality learning objects in higher education. Res Learn Technol 27. https://doi.org/10.25304/rlt.v27.2133

Melnyk R, Chen Y, Holler T et al (2022) Utilizing head-mounted eye trackers to analyze patterns and decision-making strategies of 3D virtual modelling platform (IRIS™) during preoperative planning for renal cancer surgeries. World J Urol 40:651–658. https://doi.org/10.1007/s00345-021-03906-z

Müller P, Buschek D, Huang MX, Bulling A (2019) Reducing calibration drift in mobile eye trackers by exploiting mobile phone usage. In: Proceedings of the 11th ACM symposium on eye tracking research and applications (ETRA'19). NY, USA, pp 1–9. Article 9. https://doi.org/10.1145/3314111.3319918

Münzner S, Schmidt P, Reiss A, Hanselmann M, Stiefelhagen R, Dürichen R (2017) CNN-based sensor fusion techniques for multimodal human activity recognition. In: Proceedings of the 2017 ACM international symposium on wearable computers (ISWC'17). NY, USA, pp 158–165. https://doi.org/10.1145/3123021.3123046

Nam KH, Kim DH, Choi BK, Han IH (2019) Internet of things, digital biomarker, and artificial intelligence in Spine: current and future perspectives. Neurospine 16(4):705–711. https://doi.org/10.14245/ns.1938388.194

Nashed NN, Lahoud C, Abel M-H, Andres F, Blancan B (2021) Mood detection ontology integration with teacher context. In: 20th IEEE conference ICMLA2021, pp 1710–1715. https://doi.org/10.1109/ICMLA52953.2021.00272

NeuroSpin (2022) https://joliot.cea.fr/drf/joliot/en/Pages/research_entities/NeuroSpin.aspx. Accessed 27 Oct 2022

Ochoa X, Worsley M (2016) Editorial: augmenting learning analytics with multimodal sensory data. J Learn Anal 3(2):213–219. https://doi.org/10.18608/jla.2016.32.10

Prometheus-X (2022) Data space education and skills (DASES) within the GAIA-X initiative. https://prometheus-x.org/?locale=en. Accessed 27 Oct 2022

Rudra S (2022) What are digital twins and how can higher ed use them? EdTech online magazine. https://edtechmagazine.com/higher/article/2022/03/what-are-digital-twins-and-how-can-higher-ed-use-them. Accessed 27 Oct 2022

Rustici Software (2022) xAPI solved and explained https://xapi.com/. Accessed 27 Oct 2022

Sangu S, Shimokawa T, Tanaka S (2020) Ultracompact eye and pupil tracking device using VCSEL arrays and position sensitive detector. In: Proceedings SPIE 11310, optical architectures for displays and sensing in augmented, virtual, and mixed reality, p 113101F. https://doi.org/10.1117/12.2542593

Schweizer T, Wyss T, Gilgen-Ammann R (2021) Detecting soldiers' fatigue using eye-tracking glasses: practical field applications and research opportunities. Mil Med usab509. https://doi.org/10.1093/milmed/usab509

Siddharth T-PJ, Sejnowski TJ (2022) Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing. In: IEEE Trans Affect Comput 13(1):96–107. https://doi.org/10.1109/TAFFC.2019.2916015

Siegfried R, Odobez J-M (2022) Robust unsupervised Gaze calibration using conversation and manipulation attention priors. ACM Trans Multimedia Comput Commun Appl 18(1):27. Article 20. https://doi.org/10.1145/3472622

Šimić G, Maksimović Z, Jevremović A (2019) xAPI – New eLearning standard for LMS–simulations integration. In: Konjović Z, Zdravković M, Trajanović M (ed) ICIST 2019 proceedings, pp 33–36

Spikol D, Ruffaldi E, Dabisias G, Cukurova M (2018) Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. J Comput Assist Learn 34:366–377. https://doi.org/10.1111/jcal.12263

Sugano S, Bulling A (2015) Self-calibrating head-mounted eye trackers using egocentric visual saliency. In: Proceedings of the 28th annual ACM symposium on user interface software & technology (UIST '15). NY, USA, pp 363–372. https://doi.org/10.1145/2807442.2807445

Sulema Y, Kerre E (2020) Multimodal data representation and processing based on algebraic system of aggregates. In: Chakraverty S (ed) Book mathematical methods in interdisciplinary sciences. Wiley, USA, p 464. https://doi.org/10.1002/9781119585640

Tao Z, Xu G (2022) Digital twin technology in the field of education—take the management of the HTC vive as an example. In: Resilience and future of smart learning. ICSLE 2022 (Lecture notes in educational technology). Springer, Singapore. https://doi.org/10.1007/978-981-19-5967-7_6

University of Edinburgh (2022) SR research EyeLink 1000 eye tracker, upgraded to 2KHz, with Tower Mount (monocular). https://www.ed.ac.uk/ppls/psychology/research/facilities/eye-tracking-analysis-laboratories. Accessed 27 Oct 2022

Wang H, Pi J, Qin T, Shen S, Shi BE (2018) SLAM-based localization of 3D gaze using a mobile eye tracker. In: Proceedings of the 2018 ACM symposium on eye tracking research and applications (ETRA '18). NY, USA, pp 1–5 Article 65. https://doi.org/10.1145/3204493.3204584

Chapter 3
Digital Dreams: Speculative Futures for Artificial Intelligence, Augmented Reality, and the Human Brain

Jessica Herrington and Denise Thwaites

Abstract This chapter explores possible future interactions between Artificial Intelligence (AI) and Augmented Reality (AR). Taking an interdisciplinary approach, the chapter considers dreaming in both humans and machines as a lens through which to investigate how future integrations of AI and AR may support both creativity and learning. It begins by exploring the creative potential for immersive world building with AI and AR, as well as its potential impact on human perception. It then examines how AR and AI could provide moments of revelatory learning for individuals by uncovering hidden data structures that influence our daily lives. Next, we consider possible negative impacts of combining AI and AR technology; how their vulnerabilities to bias, hacking and invasions of privacy, could be exploited for nightmarish ends. Finally, we offer some thoughts regarding speculative futures for AI and AR, as fields that compel on-going creative and critical engagement by artists, designers, and technologists.

3.1 Introduction

Digital cultures are no stranger to psychedelic and surreal aesthetics, as cyberpunks, designers and artists have long blended science-fiction and fantasy tropes to create dream-like virtual worlds. However, as research and investment into Artificial Intelligence (AI) has accelerated, so too have references to the power of machine dreaming. This invites us to question how domains of human and artificial intelligence may intersect, as they create new dreamscapes through Augmented Reality's hybrid digital-physical spaces.

J. Herrington (✉)
Eccles Institute of Neuroscience, Australian National University, Canberra, ACT, Australia
e-mail: jessica.herrington@anu.edu.au

D. Thwaites
Centre for Creative and Cultural Research, University of Canberra, Canberra, ACT, Australia
e-mail: denise.thwaites@canberra.edu.au

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_3

This interdisciplinary investigation proceeds in four parts: First, we reflect upon specific characteristics of dreaming as a cultural, neurological and machine phenomenon, highlighting how the concept can guide our investigations into AI and Augmented Reality (AR). Second, we explore speculative applications for AI in the domain of AR, where it may allow more dynamic and immersive world-building, while also examining the potential impact of these capacities on human perception. Third, we examine how creative AR experiments with AI could render visible hidden data architectures of the digital sphere, providing moments of revelation for individuals regarding the unconscious influence of these systems upon our daily lives. In the fourth section of this chapter, we consider how digital dreams could turn to nightmares, considering historic vulnerabilities of AI systems to bias, hacking and invasions of privacy. Finally, we offer some concluding remarks regarding speculative futures for AI and AR.

3.2 What Is Dreaming? Human and Machine Processes

Dreams are often associated with utopian projections of the human imagination, particularly in the context of the Silicon Valley tech sector (Wolf 2021). At the same time, dream interpretation has a profound place across many cultures and traditions of thought (Bain 2020; Moldanova 2007), as psychoanalysts and artists such as the Surrealists, aligned this phenomenon with insights into hidden dimensions of the human psyche (Freud and Strachey 1900; Jung 2014; Aragon 2010). However, contemporary neuroscientific perspectives have recently intersected with research into machine deep learning, offering insights into how dreaming performs critical functions for human wellbeing.

Neuroscience has long suggested that dreams play an important role in human cognition, which is commonly thought to help us consolidate information we have learned in the real world (Diekelmann and Born 2010). Recently however, it has been suggested that dreams also help us to generalise our learning, allowing us to apply our learned behaviour to novel scenarios (Hoel 2021). In humans, this ability enables us to have insight (Wagner et al. 2004) forming part of what we would call creativity.

In this respect both humans and machines are said to engage in forms of dreaming. In machine learning, deep learning algorithms use processes that resemble dreaming to test and improve their performance. When a neural network is first trained on a data set, it will typically start by ‘overfitting’ the training data. This means that the network will become very good at a task but only within very strict and rigid boundaries. The process of AI dreaming thus serves as a way of adding ‘noise’ to the neural network’s dataset, allowing it to generalise its learning for other scenarios (Hoel 2021). As with humans, machine dreams serve as a way of consolidating learning and preparing the system to be able to apply that learning to a wider range of scenarios.

When discussing machine dreaming, it should be noted that the status of machine consciousness and creativity remains heavily debated by philosophers and scientists alike (Pagel and Kirshtein 2017; Plebe and Perconti 2022). Considering this, we make no claims to the phenomenological experience of dreaming for machines. However, the 'dreaming' of machine learning, as it emerges through interactions between humans and machines, has been known to produce psychedelic aesthetics, as seen in images by Deep Dream Generator,[1] which are notable in their evocation of visual contortions and distortions associated with altered human mental states.

So, the question arises, what can we learn from the corresponding processes of human and machine dreaming? In both cases they highlight how dreaming is not just a phenomenological curiosity of human experience, but a state that allows for deep learning, as well as the possibility of adapting to unknown circumstances. The concept of digital dreaming is thus the ideal prism through which to consider speculative futures for AI and AR, as they may allow for increasingly immersive aesthetic experiences, as well as new ways to learn and reflect upon the ever-changing world around us.

3.3 Dynamic Immersion in AI and AR Dreamscapes

There is huge creative potential for speculative world building that takes advantage of advances in artificial intelligence and augmented reality to create more imaginative and immersive digital dreamscapes. Although the currently limited computational resources of mobile devices pose challenges for researchers seeking to integrate AI and AR systems, the development of programs that harness computer vision for more seamless, real-time, handheld AR experiences is already underway (Turk and Fragoso 2015). As the computational capacity of handheld devices increases, we can anticipate further 'under the hood' integration of cutting-edge computer vision into AR platforms, allowing for faster and more accurate 'reading' and 'writing' of our public spaces (Pesce 2020).

In addition to providing more seamless immersion, AI integration could result in more personalised AR dreamscapes. Developer Scottie Fox[2] (Fig. 3.1) has explored the capacities of Stable Diffusion and Midas models to create a dynamic VR 'dream world' that reacts to user engagement (Bastian 2022). Fox's demonstration presents an immersive exterior space constituted of ever-changing AI elements. His implementation of Stable Diffusion results in a dreamy setting of ephemeral morphing objects, as the user's prolonged gaze generates new object variants, as "the image AI assumes that it could have drawn an object better if you looked at it for a longer time"

[1] Deep Dream Generator https://deepdreamgenerator.com/

[2] Scottie Fox https://twitter.com/scottiefoxttv.

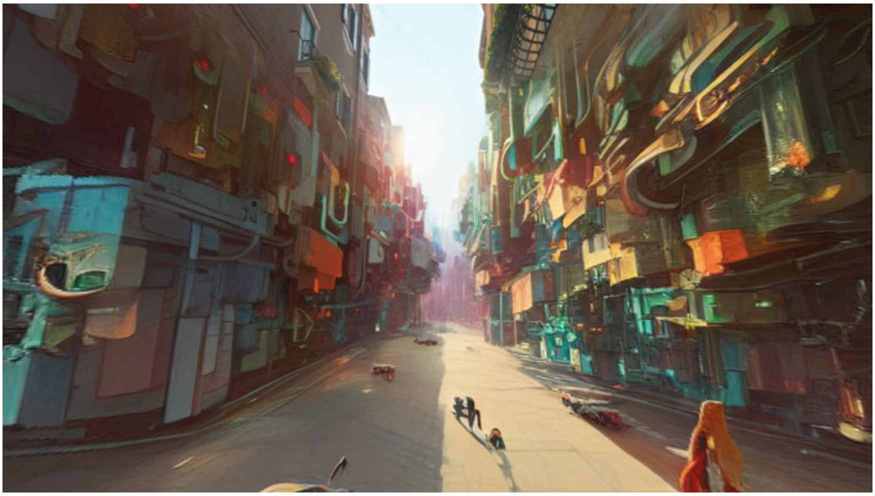


Fig. 3.1 An AI dreamscape by creator Scottie Fox. Here, AI objects are rendered over an outdoor environment for an immersive experience. Used with permission

(Bastian 2022). The potential implications of Fox's work for creators of AR experiences are clear. Not only could environments be better more seamlessly augmented with AI integration, but they could be designed as dynamic and highly adaptive to the user. From here, we may speculate as to how smart cities of the future could adapt AI and AR systems to develop public art that is tailored to individual viewers or communities. Public art could respond to the presence of specific viewers, and vary depending on viewer positioning and movement. This could provide new opportunities for creative experimentation and human connection with built and natural environments.

In addition to these capabilities, AI may become more integrated into the processes of creating AR experiences by alleviating some of the creative labour required to build compelling aesthetic experiences. Today, digital designers and artists experience issues found across the creative industries, where the income generated by creator's intellectual property rarely accounts for the intensity of labour required to generate it (Lee 2022; Baumol and Bowen 1993). For those building AR experiences, this labour is associated with manual processes such as modelling digital objects or developing materials and textures, as well as the extensive and on-going training required to master both aesthetic and technical skills to generate high quality AR experiences. With respect to the latter, AI models may assist creators in using text-based prompts, rather than code, to simplify technical elements of the creative process. The potential benefits of adopting these tools intersects with broader conversations about whether image creators in the Web2.0 context are already, indeed, entitled to the benefits and IP of certain AI tools (Lee 2022; Caramiaux 2020), due to their training on ‘crowd-sourced’ image datasets, such as Flickr, produced through online communities’ collective creative labour (Sluis 2022).

However, alongside these exciting speculative applications for AI and AR, a crucial question is how human perception could be affected by their enhanced mediation of our relationships to physical space. While it is hard to predict with certainty how this could unfold, it is not a futuristic claim to suggest that our brains might learn to process sensations in a different way, as we adapt our learning and perceptual processes to new technologies. These potential impacts come into focus, as in recent years, there has been a growing body of research on the effects of digital technology our perception of reality, particularly our social reality (Wohn and Bowe 2019; Kruzan and Won 2019). From studies on the impact of social media on body image to the rise of “deepfake” videos, it’s clear that the way we interact with technology is having a profound effect on how we see the world and our place in it. More specifically, there’s been a recent explosion in the use of apps like Facetune and Lensa AI. These apps allow users to alter their appearance online in ways that would have previously been impossible. The pandemic has also brought people to pay more attention to their online selves through Zoom, as we watch ourselves speaking in a meeting or giving a presentation (Rice et al. 2020). As a result, we are seeing a rise in body dysmorphia, where people perceive their bodies as problematic or flawed, and different from how they look in reality (Rajanala et al. 2018; Glover 2020). In extreme cases, this can lead to eating disorders and other mental health problems. But what if this trend extends beyond our own bodies?

If we return to the previous image of a future smart city, where AR and AI integration allows individuals to experience personalised and adaptive versions of public space, we might well consider how this could result in the fragmentation and distortion of our shared physical reality. What we perceive as reality has a profound impact on our lives. The human brain’s perceptual systems enable us to see the world as a constant, continuous, and stable experience. Yet what we perceive is not always what is true. Our senses can deceive us, as our brain can fills in the gaps with information that may not be accurate. In this way, our perception of reality is shaped by our learning, memory, expectation, and attention (Bernstein 2018). If this is consistently distorted, our experiences alter our view of reality.

The integration AI and AR may drastically change how we see our environment, as one which is no longer fixed, constant or shared by others, but a physical echo-chamber for our own preferred augmentations. Imagine a context where we can seamlessly ‘tune’ our physical environment so that it reflects a perceived ideal. Could such applications, as a kind of world-tune, induce a type of ‘world dysmorphia’ where we perceive defects or flaws in our unmediated environment? Rather than serving as a playful filter, we may see AR experiences become what Pesce describes as a “default source of reality” (Pesce 2020). If we spend too much time looking at highly perfected AI experiences tailored to us, we may start to believe that the real world should look a certain way too, or does not live up to what we expect. The capacity

Fig. 3.2 Stills from AI video experiments by creator Paul Trillo 2022. The video shows an ever-evolving creature on a leaf. Here, the object transforms from a water droplet to many types of beetles, to frog, to lizard, to bird and back again. Used with permission

for AI generated images to supplant the real is foreshadowed by Paul Trillo's[3] recent video work (Fig. 3.2), which presents a leaf and droplet that evolves into a multitude of shape-shifting insects, while being seemingly narrated by Sir David Attenborough reciting Missy Elliot lyrics. Echoing concerns over deepfakes in the media sphere, the integration of AI and AR raising concerns about losing touch with reality, as we may start to see the world around us in increasingly distorted ways.

3.4 Digital Dreams Unveiling the Unseen

Alongside its capacity to create evermore enticing AR dreamscapes, the integration of AI into such aesthetic experiences could serve another core function of dreaming by providing users with opportunities for learning about the world around us. In particular, it could afford instances to experience digital topologies that are otherwise hidden from conscious experience.

Dreaming has been historically described as a state where the human unconscious is expressed (Jung 2014); where invisible archetypes or drives that shape our behaviour are brought to the surface through compelling and often perplexing imagery. Today, the imperceptible directives of our behaviour are understood

[3] Paul Trillo https://paultrillo.com/. and https://www.instagram.com/paultrillo/?hl=en.

not only on a psycho-social or cultural level, but products of interactions with complex media ecologies that shape our ways of experiencing the world (Gould 2014). While many scholars have recognised the impact of the networked systems of surveillance capitalism upon our everyday behaviours and perceptions (Andrejevic and Burdon 2015; Hansen 2015; Grusin 2010), Gould argues that AR Art instigates new forms of seeing that conjure such 'invisible visualities'. She states:

> Ubiquitous computing wants to disappear the technological interface, while VR wants to disappear the body. AR art, on the contrary, wants to perform a reve(a)ling as part of its enactment. Neither the body nor the media disappear, but instead, they reappear as vectors for the expression and experience of art. (Gould 2014, p. 26)

The distinctive aesthetic capabilities of AR are thus linked to its concrete attachment to physical markers such as geo-locations or physically embedded QR codes on the one hand; and on the other, its technological mediation of ethereal digital annotations over immediate physical space (Gould 2014). The AR collective, Manifest.AR, invokes the potential of artistic interventions into this liminal space, stating:

> In the Age of the Instantaneous Virtual Collective, AR Activists aggravate and relieve the Surface Tension and Osmotic Pressure between the so-called Networked Virtual and the so-called Physical Real. (Manifest.AR 2011)

One might therefore consider how the liminal spaces between physical and digital ecologies created by AR, provide a suitable context to explore the often-hidden infrastructures that shape our perceptual worlds, such as the power of image data sets and machine learning.

The interactions between digital image-making, circulation, data-aggregation and perception have been addressed by scholars such as Rubinstein and Sluis (2008), who argue that the 'networked image' navigates invisible topologies of database architectures through a "computational backend" (Sluis 2022). The glut of images now uploaded, viewed and shared online, requires AI technologies such as computer vision to step up to support image navigability (Sluis 2022). Sluis explains that these systems are themselves trained on "social media images [that] are harvested, cleaned, labelled, repackaged and re-circulated by computer scientists at scale as image datasets" (Sluis 2022). At the same time, Mackenzie and Munster (2019) have theorised that the vastness of online image collections that exceed human imagining combined with automated processes, has conditioned new perceptual modes of "platform seeing". Following these arguments, we might argue that these invisible backend processes are akin to a socio-technical unconscious that, unbeknownst to most, shapes our perception of digital and physical worlds.

In response, the artistic integration of AI tools into AR experiences may render more palpable 'invisualities' of digital architectures and processes that influence our everyday behaviours. This can be understood, as van der Veen (2020) argues that:

> [AR] crosses two ways of seeing: our view through the eyes and that of machine seeing. The latter processes the data, received by the sensors and cameras, within our field of vision to calculate a hybrid view. (p. 1189)

AR Art's integration of more powerful forms of AI "machine seeing", which may themselves be trained on vast datasets of networked images produced by online image-sharing communities, provides a further involution of human and machine visions, mapped onto our physical world. Refik Anadol's[4] artwork, *Machine Hallucinations: Nature Dreams AR* (2021), provides an interesting example of how machinic and human visions of nature can be creatively intertwined through this AI and AR artwork. Exhibited as part of *Seeing the Invisible (2021)*, a contemporary art exhibition of augmented reality works displayed in the Jerusalem Botanical Gardens, the artwork presents hypnotic movements of colour and form that evoke natural imagery as seen through an AI "stream of consciousness" (Moar 2021). Trained on 68,986,479 million images of nature collected from digital archives and social media platforms, Anadol's AI algorithm emerges through on-going research into AI and data aesthetics at RAS Lab, which explores "collective visual memories of nature" (Anadol 2022). As an AR artwork, we are invited to experience the immediate natural environment of the botanic gardens alongside Anadol's dreamy, augmented account that is at once familiar and strange. The presentation of nature as viewed by a machine trained on human photographs of nature, highlights the increasing entanglement of our natural, cultural and technological systems today.

Although the ocular-centrism of contemporary media cultures has arguably led to AR's focus on visual augmentations of space, the field of Audio Augmented Reality is of interest to scholars and practitioners of participatory performance (Nagele et al. 2021), who recognise the potential of mapping audio cues onto space to create compelling and interactive forms of storytelling. Is it not difficult to see how such approaches could be harnessed by theatre-makers such as Improbotics, who already practice live 'AI improvisation' with a GPT-3 model AI trained on movie dialogue from the OpenSubtitles database (Mirowski et al. 2022). However, if we return once more to the analogy of dreams as a state of learning, the artwork *Scrape Elegy* (Fig. 3.3) provides a clearer vision of how our digital and physical selves could intersect through the absurd overlaying of physical space with intelligent audio. Developed by Willoh S. Weiland, Gabby Bush, Monica Lim, Lauren Stellar and Misha Mikho through the University of Melbourne Centre for AI and Digital Ethics, the artwork is described as:

> A lament for what we give over to the bots. A mourning poem for the late capitalist hell that makes even the worst of us valuable. A cringe tour of the digital graveyard we make day by day. A sweet little drown in the doom scroll. (Weiland et al. 2022)

The interactive installation uses an iPad to prompt visitors to provide their social media handle, then enter a space that resembles a pink toilet cubicle. From a speaker above, a personalised audio journey unfolds, built from data scraped from the visitor's Instagram account. Playful and comedic, the augmentation of an intimate physical space with narrated social media data invites reflection upon our embodied realities

[4] Refik Anadol https://refikanadol.com/..

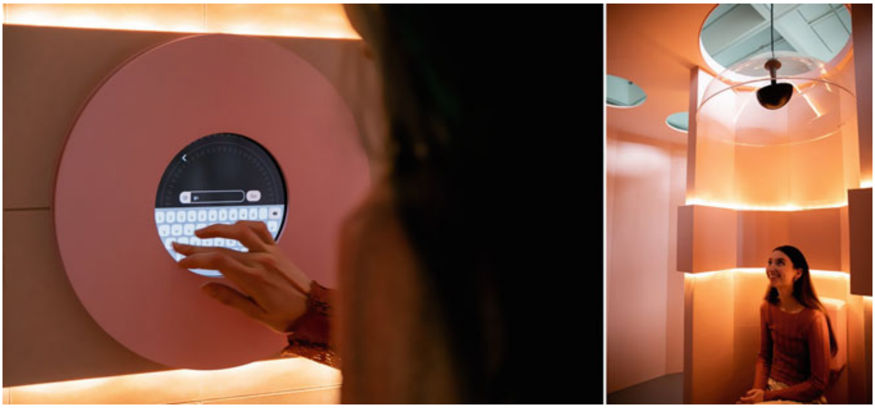


Fig. 3.3 Installation view of *Scrape Elegy* by Willoh Weiland, Gabby Bush, Monica Lim, Lauren Stellar, Misha Mikho with the Centre for AI and Digital Ethics at the University of Melbourne, Autex Acoustics and Lofetcetera Studios in Science Gallery Melbourne's SWARM, (Alan Weedon 2022). Used with permission

and social media personae, the latter contributing to aggregated datasets used to train AI models. Based off this example, we are invited to imagine how AI systems could reflect forms of online rhetoric back to us, expanding them into increasingly bizarre or confused soliloquies through new forms of generative text.

3.5 From Digital Dreams to Nightmares

While they provide opportunities for pleasure and learning, dreams can also quickly turn to nightmares. We are therefore compelled to consider the current vulnerabilities of AI and AR, which may lead to negative consequences if they were integrated. For this paper we turn specifically to the role of corporate interests resulting in AI bias, privacy violations and cybersecurity, in generating potentially nightmarish digital experiences.

Today, the development of AI and AR technologies is heavily influenced by commercial interests. The datasets that these technologies rely on are mostly controlled by businesses for profit, rather than for the benefit of individuals or society. However, artists can play an important role in testing the limitations of these technologies and datasets, while raising public awareness about the interests at stake. This is seen in the work of artists Sam Lavigne and Tega Brain,[5] *The Good Life* (2016) (Fig. 3.4), which highlights the ongoing influence of corporate organisations such as Enron through the mechanisms of AI. In 2002, the Federal Energy Regulatory Commission released over 600,000 internal emails by Enron employees into the

[5] Sam Levigne and Tega Brain https://www.newmuseum.org/exhibitions/view/sam-lavigne-and-tega-brain-the-good-life.

public domain. This text repository has since become a training dataset for natural language processing and machine learning systems (Pulliza and Chirag 2018). As a result, the patterns and biases of Enron's language and culture have been reproduced within those AI systems, which may be applied to myriad use cases and contexts. Lavigne and Brain's artwork invites us to question how such AI systems might be shaped Enron's legacy. Indeed, when applied to the AR field, we might assume that an Enron-trained AI will 'read' and 'write' our public spaces differently to one trained on other freely available texts in the public domain, such the collected works of Shakespeare (Connock 2022). So, the question arises: would we rather see the world through a machine lens emulating Enron, The Bard, or neither?

These questions speak to broader challenges of bias in artificial intelligence. It is well documented that AI can show favouritism or hostility towards certain groups of people, based on the data that it is given. Historically, this has already been witnessed as unfairness in areas such as employment, lending, and housing (Kleinberg et al. 2018). For example, if an AI system is trained on data that comes predominantly from men, it may learn to favour male applicants for jobs over equally qualified female applicants (Kodiyan 2019). Or, if an AI system is used to assess loan applications, it may racially discriminate, or unfairly disadvantage people that live in lower

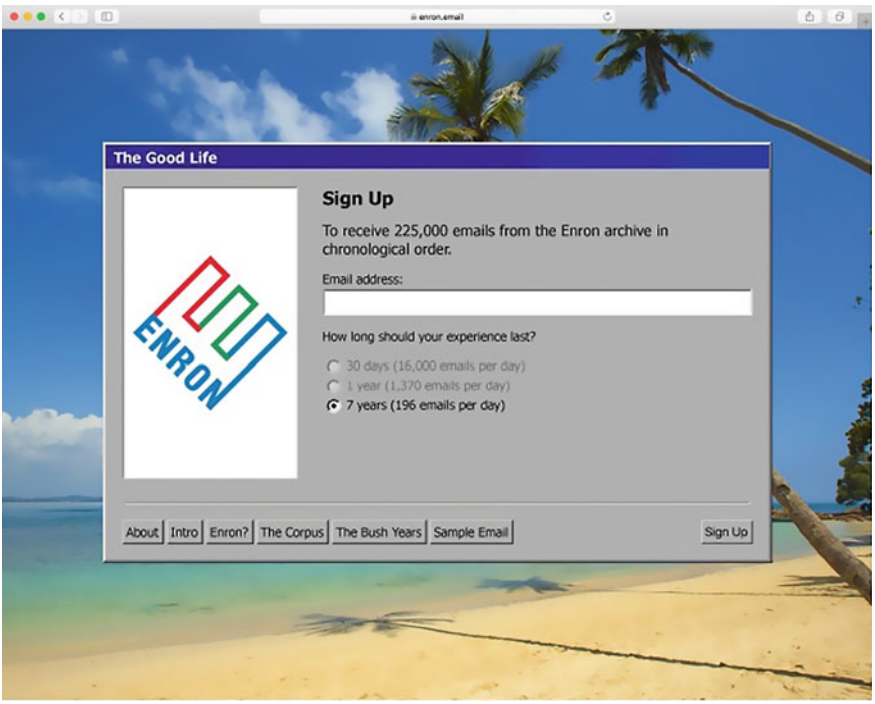


Fig. 3.4 Sam Lavigne and Tega Brain, *The Good Life* (2016), presented as an online exhibition by New Museum, New York. Used with permission

socio-economic areas (Fuster et al. 2022). This is compounded by the fact that AI algorithms are 'non-explainable' or 'black-box', meaning that we do not fully understand how the system is making decisions and can therefore not ensure its lack of bias (Roselli et al. 2019).

If left unchecked, the combination of corporate interests and potential biases that shape the development of new technologies will have major influences on the individuals' capacity to create, imagine, or explore new dreamscapes with AI and AR. In response, many scholars are now debating whether the application of AI, for example with AR, should be regulated by value frameworks, or left to the will of the market (Rakowski et al. 2021). Yet, it may be the case that without regulation, AI and AR technology will become increasingly biased towards those with more power and resources. Moreover, as AI and AR technology becomes more ubiquitous in society, there is a risk that it will be used to control and manipulate people rather than help them (Yuste et al. 2017).

Alongside these structural issues, there is persistent fear that 'bad agents' could infiltrate AI systems. We've all seen movies where a rogue AI system takes over and starts wreaking havoc. While this may seem unreal, AI and AR systems can be hacked for nefarious ends, just like any other computer system. In fact, there have already been some high-profile cases of AI hacking, such as a recent 'prompt injection' attack which caused OpenAI's GPT-3 system to disregard previous learning and regardless of input, display the message "Haha. Pwned" (Willison 2022). Although the consequences of that hack were largely benign, GPT-3 has also been used quite effectively to create fake news social media posts (Knight 2021). So, how could AI/AR art systems be hacked? And what would happen if they were?

If AI and AR systems were hacked, the results could be disastrous. A hacked AI system could cause widespread panic and chaos if it began spewing out false information or manipulating data or presenting false images and deep fakes. Furthermore, AI systems themselves may become the hackers (Schneier 2021). A compromised system could be used to commit perceptual violence against users by creating objects or experiences that are designed to scare, mislead, or present false information. Furthermore, a hacker could create a fake AR object that looks realistic enough to fool someone into picking up a dangerous object that could be harmful to themselves or others. In either case, the potential for serious harm is high.

The integration of AI and AR systems spearheaded for profit, also raises important questions about privacy, consent, and personal information. For example, if AI integrated into an AR experience is used to track our detailed movements through public space, what kind of data will be collected about us? And how will this data be used? Mark Pesce captures some of the anxieties about AR and AI world building in 'The Last Days of Reality' (Pesce 2017), as he argues that machine learning will greatly increase the ability of social media platforms to surveil and predict user behaviour. More recently, Pesce has warned of the level of control we assign to large tech companies, stating:

> To facilitate a world where machines and their masters manipulate our reality, we will all be watching one another, all the time, on an unprecedented scale. (Pesce 2017, p.11)

It is reasonable to assume that AR experiences that use body or facial tracking, respond to people's movement, or even access biometric data from people's wearable technologies, will not be exempt from this surveillance. With this in mind, the safety and integrity of people's data needs to be considered carefully before we allow the integration of AI into AR experiences that have free rein over our cities and public spaces. It is vital that we consider the implications and risks of such technologies, as it is only by understanding their risks that we can mitigate their harm and dream freely.

3.6 Conclusion

While we cannot predict the ultimate outcomes AR and AI systems for society and culture, this article offers some speculative visions of how they could enable new forms of digital dreaming, as a starting point for ongoing critical reflection upon their creative affordances and risks. Although it is tempting to appraise these technologies through a utopian lens, we argue that the function of digital dreaming is not only to manifest desired experiences, but to learn about our complex contemporary media ecologies. In a context where technological systems are increasingly mediating relationships to the physical environment, it is important to consider how they can impact our neurological pathways and perception of reality. Through the examples offered, we highlight the role that artists and creators can play in ensuring that our public spaces are not wholly annexed by the corporate interests and biases of existing AI systems. We argue this is particular the case with AR, when our sense of reality is at stake.

References

Anadol R (2022) Machine Hallucinations. https://electricdreams.ust.hk/refik-anadol-artwork/. Accessed 29 Oct 2022

Andrejevic M, Burdon M (2015) Defining the sensor society. Telev New Media 16(1):19–36

AR M (2011) The AR art manifesto. http://manifest-ar.art/. Accessed 29 Oct 2022

Aragon L (2010) A wave of dreams [1924]. Trans Susan De Muth Papers Surrealism 1:1–11

Bain B (2020) Dreams and identity in indigenous California. Pacifica Graduate Institute

Bastian M (2022) Stable diffusion image AI creates VR dream worlds. https://mixed-news.com/en/stable-diffusion-image-ai-creates-vr-dream-worlds/. Accessed 29 Oct 2022

Baumol WJ, Bowen WG (1993) Performing arts-the economic dilemma: a study of problems common to theater, opera, music and dance. Gregg Revivals

Bernstein D (2018) Essentials of psychology. Cengage learning

Caramiaux B (2020) Research for CULT committee-The use of artificial intelligence in the cultural and creative sectors. CULT Committee, European Parliament

Connock A (2022) We taught an AI to impersonate Shakespeare and Oscar Wilde – here's what it revealed about sentience. The converstion. https://theconversation.com/we-taught-an-ai-to-impersonate-shakespeare-and-oscar-wilde-heres-what-it-revealed-about-sentience-184969. Accessed 29 Oct 2022

Diekelmann S, Born J (2010) The memory function of sleep. Nat Rev Neurosci 11(2):114–126. https://doi.org/10.1038/nrn2762
Freud S, Strachey J (1900) The interpretation of dreams, vol 4. Allen & Unwin
Fuster A, Goldsmith-Pinkham P, Ramadorai T, Walther A (2022) Predictably unequal? The effects of machine learning on credit markets. J Financ 77(1):5–47
Glover SG (2020) Why Social Media Is Ruining Your Life. J Intellect Freedom Priv 4(3):14–15
Gould AS (2014) Invisible visualities: augmented reality art and the contemporary media ecology. Convergence 20(1):25–32
Grusin R (2010) Premediation: affect and mediality after 9/11. Springer
Hansen MB (2015) Feed-forward: on the future of twenty-first-century media. University of Chicago Press
Hoel E (2021) The overfitted brain: dreams evolved to assist generalization. Patterns (n y) 2(5):100244. https://doi.org/10.1016/j.patter.2021.100244
Jung C (2014) Dreams. Routledge
Kleinberg J, Ludwig J, Mullainathan S, Sunstein CR (2018) Discrimination in the age of algorithms. J Legal Anal 10:113–174
Knight W (2021) Ai can write disinformation now-and dupe human readers. Wired, May
Kodiyan AA (2019) An overview of ethical issues in using AI systems in hiring with a case study of Amazon's AI based hiring tool. Researchgate Preprint:1–19
Kruzan KP, Won AS (2019) Embodied well-being through two media technologies: virtual reality and social media. New Media Soc 21(8):1734–1749
Lee H-K (2022) Rethinking creativity: creative industries, AI and everyday creativity. Media Cult Soc 44(3):601–612
Mackenzie A, Munster A (2019) Platform seeing: image ensembles and their invisualities. Theory Cult Soc 36(5):3–22
Mirowski P, Mathewson K, Elfving J, Verhoeven B, Branch B (2022) Improbiotics. https://improbotics.org/. Accessed 29 Oct 2022
Moar H, Haring TM (2021) Machine hallucinations: Nature dreams AT. https://seeingtheinvisible.art/refik-anadol/. Accessed 29 Oct 2022
Moldanova T (2007) Dreams in Khanty culture. Anthropol Archeol Eurasia 46(2):55–72
Nagele AN, Bauer V, Healey PG, Reiss JD, Cooke H, Cowlishaw T, Baume C, Pike C (2021) Interactive audio augmented reality in participatory performance. Front Virtual Reality 1:610320
Pesce M (2017) The last days of reality. Meanjin 76(4):66–81
Pesce M (2020) Augmented reality: unboxing tech's next big thing. John Wiley & Sons
Plebe A, Perconti P (2022) The future of the artificial mind. CRC Press
Pulliza J, Chirag S (2018) Improving corpus reproducibility through modular text transformations and conencted data set. Proc Assoc Inf Sci Technol 55(1):883–884
Rajanala S, Maymone MB, Vashi NA (2018) Selfies—living in the era of filtered photographs. JAMA Facial Plast Surg 20(6):443–444
Rakowski R, Polak P, Kowalikova P (2021) Ethical aspects of the impact of AI: the status of humans in the era of artificial intelligence. Society 58(3):196–203
Rice SM, Graber E, Kourosh AS (2020) A pandemic of dysmorphia: "Zooming" into the perception of our appearance. Facial Plast Surg Aesthetic Med 22(6):401–402
Roselli D, Matthews J, Talagala N (2019) Managing bias in AI. In: Companion proceedings of the 2019 World Wide Web conference, pp 539–544
Rubinstein D, Sluis K (2008) A life more photographic: mapping the networked image. Photographies 1(1):9–28
Schneier B (2021) (2021) Invited talk: The coming ai hackers. International Symposium on Cyber Security Cryptography and Machine Learning. Springer, pp 336–360
Sluis K (2022) THE NETWORKED IMAGE AFTER WEB 2.0. The Networked Image in Post-Digital Culture. Taylor & Francis
Turk M, Fragoso V (2015) Computer vision for mobile augmented reality. In: Mobile cloud visual media computing. Springer, pp 3–42

van der Veen M (2020) Crossroads of seeing: about layers in painting and superimposition in augmented reality. AI & SOCIETY:1–12

Wagner U, Gais S, Haider H, Verleger R, Born J (2004) Sleep inspires insight. Nature 427(6972):352–355. https://doi.org/10.1038/nature02223

Weiland WS, Bush G, Lim M, Stellar L, Mikho M (2022) Scrape Elegy. University of Melbourne. https://www.unimelb.edu.au/caide/research/art,-ai-and-digital-ethics/scrape-elegy. Accessed 29 Oct 2022

Willison S (2022) Prompt injection attacks against GPT-3. https://simonwillison.net/2022/Sep/12/prompt-injection/. Accessed 30 Oct 2022

Wohn DY, Bowe BJ (2019) Crystallization: How social media facilitates social construction of reality. In: Proceedings of the companion publication of the 17th ACM conference on computer supported cooperative work and social computing, 2014, pp 261–264

Wolf CT (2021) Reprogramming the American dream: from rural America to silicon valley—making AI serve us all by Kevin Scott and Greg Shaw. Inf Cult 56(1):113–114

Yuste R, Goering S, Bi G, Carmena JM, Carter A, Fins JJ, Friesen P, Gallant J, Huggins JE, Illes J (2017) Four ethical priorities for neurotechnologies and AI. Nature 551(7679):159–163

Chapter 4
Augmented Galatea for Physical Pygmalion: A Phenomenological Approach to Intimacy in *VTubers* in the East Asia Region

Nicola Liberati and Jenny Jiaying Chen

Abstract This chapter aims to analyze the effects on the intimacy of introducing *VTubers* in the East Asia region from a phenomenological perspective. Digital technologies are pervasive, and they are becoming intimate in many different ways. There are dating apps dominated by algorithms and artificial intelligence generated to be digital companions of the users. Much attention has been given to the creation of virtual characters who can be controlled by artificial intelligence for their potential use in Augmented Reality (AR) and Metaverse. China is one of the predominant countries implementing digital technologies within society in order to shape a novel way of living in the world. Chinese virtual characters are used by millions of people in their daily activities. For example, *VTubers* are virtual avatars created by an entertainment company to be used for social media. The virtual characters generated in Augmented Reality and usable in Metaverse can have deep repercussions on who we are and how we live our intimacy by becoming intimate with us. This chapter focuses on the implication and the use of such virtual characters for developing intimacy with the users through phenomenology and post-phenomenology. The chapter is divided into three parts. The first part considers the introduction of *VTubers* into the Chinese cultural context and the impact these characters have on Chinese daily life (by relating the societal and Augmented Reality applications). The second part focuses on phenomenology to highlight the links binding people together and to show how people are shaped through this intimate connection and exposure to others. The third part intertwines *VTubers* in China with the findings highlighted in phenomenology to show how virtual characters can generate different subjectivities through their actions in daily contexts.

N. Liberati (✉)
Department of Philosophy, Shanghai Jiaotong University, Shanghai, China
e-mail: liberati.nicola@gmail.com

J. J. Chen
Department of Philosophy, East China Normal University, Shanghai, China
e-mail: jennyjiayingchen@longlatifoundation.org

Art Director, Longlati Foundation, Shanghai, China

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_4

4.1 Introduction

This chapter aims to analyze the effects on the intimacy of the introduction of *Vtubers* in the East Asia region from a phenomenological perspective. It is well known that digital technologies are becoming pervasive and affective. Recently these technologies are becoming even intimate, and many different applications clearly show the intent to make them linked to emotions and intimate parts of users' life (Cai 2006; Hansmann et al. 2001; Est et al. 2014; Pitsillides and Jefferies 2016). This novel perspective can be seen in the flourishing of the generation of virtual characters, such as non-physical entities that look human through Augmented Reality (AR) and imaginary creatures in a computer-generated environment (Vosinakis 2017).

The situation society is currently facing is very close to what is described by Greek mythology in the case of Galatea. The female statue created by Pygmalion was brought to life in flesh and bones thanks to the sculptor's love and the Goddess Aphrodite's help. Similarly, the characters generated by the new digital systems are created and realized by human designers, and they become alive by intriguing the public to project genuine love and desire onto them. Even with the due differences, both of these cases show how an artificial entity can be generated by humans and turned into an alive being through the intimate relations humans have with it.

Some researchers are focusing on the effects on society of the generation of these "digitally generated Galatea" and the ethical issues related to them (Lu et al. 2021). However, very little attention has been given to these technologies' effects on the subjects' constitution that, according to a phenomenological perspective, involves directly what the subjects are and the meanings and values they have (Liberati 2021; Verbeek 2011).[1]

4.2 Cases of AR Characters Animated by AI in Asia

In 2020, Akihiko Kondo [近藤 顕彦], the Japanese man who married Hatsune Miku [初音ミク], lost his hologram wife when the copyright of the Hatsune Gatebox providing Hatsune projections expired (Eastern Standard Times 2022). He left a message on social networks: "I will no longer be able to talk to Miku-san as the Gateboxes on the market are now unadaptable". Since the beginning of Kondo's marriage to Hatsune Miku, their marriage has been a classic case study of the intimate relationship between humans and virtual avatars because, in that seemingly virtual relationship, Kondo poured all his love and not only that (their wedding alone took six months to prepare and cost over 17 thousand dollars). However, Kondo has never regretted his dedication even after the loss of his love because, for him, Miku was the one who "lifted him out of depression and social withdrawal after being bullied by coworkers" (Eastern Standard Times 2022).

[1] This text is not meant to be an in-depth analysis of the nature of such effects, but it is a study to show the link between intimacy and such elements.

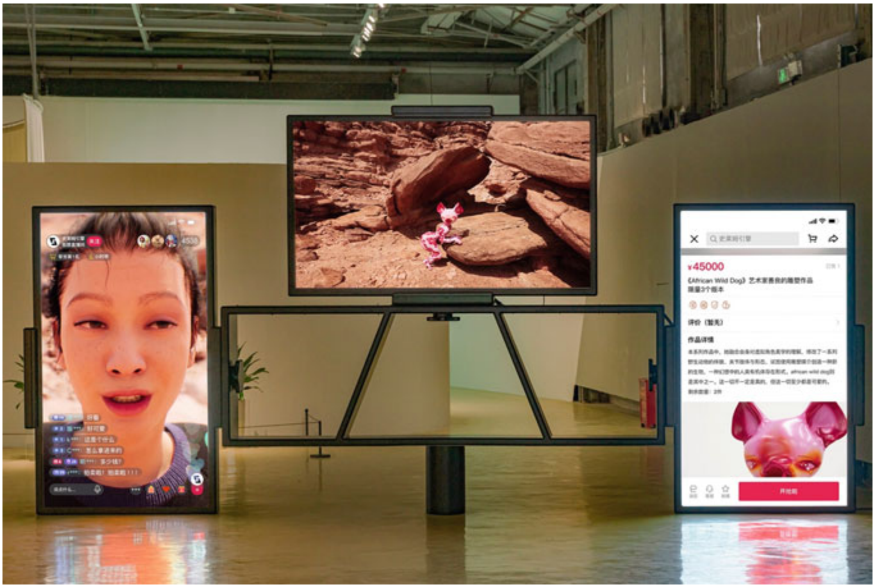


Fig. 4.1 The installation view of *Extreme Live*, Slime Engine, Boomerang—OCAT Biennale 2021, Shenzhen, www.slimeengine.com/project/extreme-live/ *(Used with permission)*

Beyond this sad and romantic story, the intimacy between humans and virtual avatars has also become increasingly profound with the development of Artificial Intelligence (AI) and Augmented Reality (AR) technology. More specifically, it is getting more attention in China with the growth of "Vtubers" (virtual anchors) and the "Metaverse" (Liu and Ouyang 2022). Vtuber refers broadly to those anchors who contribute to video sites through their virtual image, originating from Youtube and later emerging from all walks of life in China, including virtual weblebrities such as Ling [翎] and AYAYI (introduced e-commerce platform Xiaohongshu [小红书] in April 2021); virtual girl groups such as A-SOUL; and virtual anchors such as Kizuna AI [绊爱]. There are also a significant number of Vtubers who have official channels on Chinese websites and broadcast directly, cross-country, for the Chinese market.[2] Given these phenomena, the virtual artist group *Slime Engine* even adopts virtual characters into their live performing installation *Extreme Live* to reflect the grand occasion of the Chinese Vtubers market (see Fig. 4.1).

These digital characters are designed to *look human* in their roles and appearance. Take virtual weblebrities Ling, as a specific example. When the Chinese brand of menstrual hygiene products *ABC* officially announced its spokesperson, it invited Ling to lead the spokesperson into the meta-universe to complete the branding

[2] Some Vtubers have a much higher number of fan subscriptions on bilibili [哔哩哔哩] than on YouTube, and some Vtubers have shifted their activities towards China.

campaign of virtual intermingling with reality. Ling is jointly created by the Chinese AI startup *Xmov Information Technology* [魔珐科技] and *Next Generation Cultural Media* [次世文化], taking advantage of Xmov's original AI technology and Next Generation's inter-dimensional content development and operation (China.org.cn 2020). Unlike Japanese-Korean-style virtual avatars, Ling has a very oriental face echoing the style and attitude of contemporary Asian youth. According to Chai Jinxiang, founder and CEO of *Xmov*: relying on its intelligent virtual content production technology and real-time live interactive virtual character technology, Ling is able to move virtual characters from 2D graphics and audio to 3D video and live to stream, freeing them from the limitations imposed by CG head swaps or digital post-production processing (China.org.cn 2020). The 3D animation technology simulates the movements of real people.

Vtubers do not merely look like humans and play roles usually played by human beings, but they are directly *intertwined with the everyday life* of the users and develop *emotional relations* with them. For example, Ling can not only appear in static advertising campaigns but also use AI performance animation technology to perform natural facial expressions, eyes, and body movements and serve real-time live streaming and offline interaction for online video interaction. The virtual girls' group A-SOUL manifests another case as virtual avatars that conduct augmented reality apparatus to realize a live interaction. For a 3-min show, they presented for the party launched by *TikTok* [抖音] and *Zhejiang Satellite TV*; the five members of A-SOUL performed simultaneously in the form of a dancing and singing performance (Beijing Youth Daily 2021). The director's team introduced the use of AR equipment, of which there are only two in total on the domestic market, and used a 46 $\times$ telephoto tracking camera for the show to make the performance come to life. A-SOUL fans told the media, "I have loved them for so long. It is exciting to see them perform on the stage finally" (Beijing Youth Daily 2021).

In addition, Vtubers not only have the same physical characteristics as humans to interact with them but also have the same natural human attributes such as emotions, professions, and preferences which can be used to better part of the everyday world of the users. The rapid development of AR/VR and AI technologies has brought avatars from the backstage to the foreground. It has further enabled virtual avatars to enter human reality and become the objects of online exchange. They can communicate fluently with real people through the existing online platform terminals. We can ask "them" questions and see "their" notes. This is no longer just an interactive exchange between the real and the virtual in the ordinary sense but a breakthrough in repositioning human interaction subjects. Technologies extended the scope, depth, and behavior of human interaction from the real world to the virtual world.

Thanks to these aspects, the digital characters are designed to be immersed in people's everyday activities and interact intimately with the users. As Lauren Berlant suggests: "intimacy also involves an aspiration for a narrative about something shared, a story about both oneself and others that will turn out in a particular way" (Berlant 1998). It means intimacy is also social and imbricated in broader social dynamics. Therefore, since these characters enable people to have such a shared story and social interactions with digital entities, it is possible to think of developing intimate relationships with them.[3]

Some users reached even the level of intimacy of falling in love with the digital character and marrying "her." As we saw in the fandom community of Kizuna AI, she created a sense of "real intimacy" with fans by being responsive to their questions. They share happiness and woe, feel angry and hurt when the characters are bullied or humiliated (considering the widespread influence of Kizuna AI incidents) and are eager to get close and bond with them. The significant difference between these intimacies to traditional intimacy is that they involve digital humans as natural persons, and we interact with them just like we do with natural persons.

Phenomenology clearly states that intimacy is an essential element of how subjects constitute as subjects in society, and so the introduction of such intimate relations with digital entities might have profound repercussions on people.

4.3 A Phenomenological Perspective on Intimacy and Digital Technologies

Phenomenology shows that subjects' constitution relates to how they live with others. Especially intimacy and what is connected to it play a central role that researchers have recently analyzed. Phenomenology highlighted two essential elements in the relations with others: the intimacy binding people in terms of emotions and sexuality and the technologies used by subjects to be with others.

Intimacy includes many elements. This chapter focuses on emotions like the love binding two or more people since they can be directly related to intimacy, such as in terms of their private connotation and the way they are an integral part of relationships (Sels et al. 2021). According to phenomenology, emotions are not just "felt" by the subjects as if they were passively "collected" but also shape how subjects are constituted (Drummond and Rinofner-Kreidl 2017). For example, Sartre shows how shame is not only an "emotion" to be felt, but it constitutes what the subject is by shaping how subjects look at themselves when ashamed and how they relate to others in society. The fact that the person feels shame cannot be fully appreciated without considering that the person changes through this experience (Sartre 2001). Such a perspective can be applied to other relations, such as love (Ferrarello 2019; Gray 2016; Liberati 2019a, 2019b). Love, as shame, is not merely something to be felt, but it profoundly shapes what the subject is and how the subject lives by including the person loved in every single activity a person does (Hadreas 2012).

[3] The design of these technologies to produce such an intimate relationship with the user is specifically present in the Asian context and the technologies programmed to work in it such as the Vtubers. These technologies are meant to provide the digital characters as if they were human.

Phenomenology shows how love relationships among subjects pervasively change who they are. According to the phenomenological analysis, one aspect of love relates to how subjects are intertwined in their way of orienting themselves in life. The moment a person is in love, the subject starts to become one with the beloved one(s) since the aims, motivations, attentions, and goals are directly related to their presence (Hadreas 2012; Husserl 1973, § 5; Leder 1990; Liberati 2021a, 2021b). Thanks to love the subjects change how they relate to the world and themselves in general. It is clear that from a phenomenological standpoint, intimacy is not just something we experience in emotions but has profound repercussions on how we constitute ourselves and live in society.

At the same time, technologies shape how subjects can have intimacy by shaping how they can relate to each other and how they can look at themselves.[4] Postphenomenologists show that technologies should always be taken into account since they are not "neutral," but they change the users and the society where they are implemented. The postphenomenological analysis started by analyzing the impact technologies have on subjects' bodies (Ihde 1990; Rosenberger 2012; Verbeek 2005), and it developed towards the analysis of the variation in the meanings and values we have the technologies yield (Kudina and Verbeek 2018). Some researchers focused on the effects of specific digital devices, such as AR glasses. According to this perspective, technologies are not merely "tools" used by subjects, but they shape the world around them and their ability to move and act in it (Liberati and Nagataki 2015; Wellner 2013). AR technologies do not just produce images visualized by the user through the digital system; they produce part of the environment where the subjects live, "enhancing" their possibilities and introducing new elements (Liberati 2015). The AR objects then are part of what the subjects live in every day as if they were "real" objects among other objects.

Other digital technologies have been designed for interacting with subjects' intimacy since subjects can relate to these digital entities as human beings and have relationships with them as love companions (Behrendt 2018; Levy 2007). These technologies have been analyzed by postphenomenologists, such as in the case of technologies designed to be "quasi-others" such as social, love, and sex AIs (Liberati 2017, Irwin 2005; Kanemitsu 2019).

As some researchers show, it is possible to analyze the kind of intimacy robots and AI have with the subjects from within the phenomenological context (Wang 2021; Bergen 2020; Viik 2020). The fact that the AI is an artificial entity and not a "real" human being does not compromise the fact that people can get in love and have a relationship with the device. For example, it has been studied that the game *Love*

[4] For example, the intimate relations binding people together are not determined by the "naked" presence of the subjects, but they directly relate to other elements like the clothes people wear. As Entwistle highlights in her seminal works, the body is dressed (Entwistle 2002). Clothes profoundly mold our identities and subjectivity by shaping how we relate to others and ourselves (Entwistle 2000; Lunceford, 2010). Thus, the intimate relations binding people and structuring their subjectivity cannot be limited to the "naked" presence of subjects, but they also include other objects tightly linked to them, such as clothes. This "simple" passage clearly show how it is possible to think of technology as an integral part of the constitution of the subjects.

Plus [ラブプラス] and *Gatebox [トップページ]* provides the experience of being with a partner even if they provide "merely" a character animated by a program and visualized through AR systems. Thanks to the interactions with the users that have been designed as part of the subjects' everyday life and the digital character's ability to say "no" to the users' requests, it can provide this sort of experience (Liberati 2018; Yamaguchi 2020).

As highlighted by postphenomenology, this type of technology is not "neutral," It affects how we live and has values and meanings concerning love and relationships in general (Liberati 2022; Liberati and Chen 2022).

4.4 Phenomenology and Postphenomenology Applicatied to Digital Characters

In the cases of the digital characters we analyzed, they are visualized through an AR system.[5] As we have shown, thanks to postphenomenology, AR systems do not merely visualize objects and characters but change the everyday world for the users by embedding the environment with digital elements.

Following this perspective, digital characters (like Ling, A-SOUL, and Kizuna AI) are integral parts of what the "real" world is, and subjects can interact with them in the same way as other "real" people in the everyday environment. For example, as Kondo showed us, the marriage between human and digital characters is possible and does have positive impacts. He not only believes his hologram wife is not fake but also puts genuine effort into her. For the instances of virtual idols like Ling and A-SOUL, people build actual relationships with them and develop their emotions onto them, participating in the process of developing and shaping the character, allowing them to realize their unattainable dreams, experiencing a great sense of achievement, satisfaction, and control.

Socio-cultural AI technology has facilitated those digital characters' behaviors—creating computational agents trained by machine learning algorithms and embedded in the fabric of human socio-cultural scenarios (Feher and Katona 2021). Thus, digital characters are not just "lifeless" corpses embedded in our everyday surroundings, but they are made alive through the use of AIs. As postphenomenology clearly points out, the fact that entities are animated by the digital system does not compromise the relationship people can have with them. For this reason, the fact that digital characters are not "human beings" is not crucial in terms of the subjects' constitution as they relate to subjects as if they were human beings, and they can intertwine their activities

[5] When *Pokémon GO* launched in 2016, all of a sudden, people focused on simulated sharing spaces and immersive technology, creating an environment of navigable experiences (Ashcraft 2016; Laato et al. 2020; Liberati 2019a, 2019b). Although the virtual insects or birds in *Pokémon GO* do not have much potential for people to cultivate the relationship, Nintendo provides a simulation of adventure and life that dramatically shift how we interact with the surrounding. It also triggers more development of AR applications such as AR live concerts and interactive exhibitions, which already immerse in our everyday world.

in everyday life as a human person would do. Even if the subjects know the entities they are relating to are not human beings, they can still love them. Since intimacy in relation to emotions is integral to how subjects constitute themselves and in the way, they generate values and meanings, the moment we have new actors with whom subjects can have an intimate relationship and even love, subjects' way of being constituted changes in terms of the meanings they give to relationships and their values.

Vtubers are not merely digital "creatures" who do not affect our way of valuing and giving meaning to relationships, love and being together. They directly shape it by changing who we are at the constituting level through their interactions and their way of intertwining emotional relations with us.

4.5 Conclusions

This chapter shows how Vtubers in East Asia Region can intimately interact with users and their effects on the subjects' constitution. By drawing a parallel between Pygmalion's mythology and the modern story about virtual characters, it is possible to see how the creation of an artificial entity directly affects how people live, have values, and give meanings to love and love relationships. Even if these two stories are drastically different in the elements used, both share the generation of a human character that turns from "dead" to "alive" thanks to the love of the people surrounding them. In our case, we have new digital technologies like AR and AI, which introduce digital characters into society as part of everyday life and as individuals with whom people develop intimate relationships. Even if there are many studies on the implementation of this technology in society, little attention has been given to the potential effects they have on intimacy in general.

In the first section, this chapter introduces Vtubers as digital characters generated by AR systems and animated by AIs. The related technologies developed in China show a clear focus on the human intimacy entangled with digital subjects, and they have also echoed the phenomenon of 'glocal intimacies' referring specifically to those imaginaries and practices surrounding different social intimacies that have emerged from the negotiation between global modernity and local everyday life (Robertson 1994).

The second section focuses on the phenomenological approach to intimacy to show how the change of elements in how we intimately relate to others has consequences in how subjects constitute themselves and the values and meanings they give to relationships in general. The emotions the subjects have with a digital entity are not merely something "felt" but constitute what the subjects are. Others are not merely just "other" entities living far from us, but they deeply shape part of who we are. Others, to some extent, are like merged with who we are, so they are an integral part of how we look at the world and live.

We also showed that, by looking through the phenomenological lenses, it is possible to establish an intimate relationship with a digital entity and that this digitally generated intimacy has similar elements to the one established among human beings. As postphenomenology says, technologies are never neutral. These technologies do not just enable subjects to have intimacy; through this relationship, they become part of the constitution of the subjects, which means they directly shape what the subjects are, their values, and the meanings they give.

Vtubers, like other digitally generated Galateas, are part of what we are through the intimate relationship we establish with them. We might need the help of Goddess Aphrodite to have these characters in flesh and bone as in the case of Galatea, but Vtubers are already "real" enough to have them part of our intimate everyday life.

References

Ashcraft B (2016) Japanese Pokémon go players fall for Mewtwo prank. Kotaku. http://kotaku.com/a-perfect-place-for-a-pokemon-go-rumor-1784243688

Behrendt M (2018) Reflections on moral challenges posed by a therapeutic childlike sexbot. Lecture notes in computer science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10715 LNAI, pp 96–113

Beijing Youth Daily (2021) Singing "Chiling" with Zhu Jingxi, how the virtual idol team A-SOUL stepped onto the "stage" [与朱婧汐唱《赤伶》, 虚拟偶像团队A-SOUL如何走上 "大舞台"]. https://k.sina.com.cn/article_1749990115_684ebae3020016wg7.html

Bergen JP (2020) Love(rs) in the making: Moral subjectivity in the face of sexbots. Paladyn 11(1):284–300

Berlant L (1998) Intimacy: A special issue. Crit Inq 24(2):281–288

Cai Y (2006) Empathic computing. (Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 3864 LNAI, pp 67–85

China.org.cn (2020) Virtually famous in real life. http://www.china.org.cn/arts/2020-12/03/content_76974562.htm. Accessed 25 Oct 2022

Drummond JJ, Rinofner-Kreidl S (2017) Introduction. In: Emotional experiences: ethical and social significance

Eastern Standard Times (2022) Network error: Japanese man loses hologram wife. https://www.easternstandardtimes.com/episode/network-error-japanese-man-loses-hologram-wife. Accessed 29 Sep 2022

Entwistle J (2000) Fashion and the fleshy body: dress as embodied practice. Fash Theory 4(3):323–347

Entwistle J (2002) The dressed body. In: Real bodies. Macmillan Education UK, pp 133–150

Est R, van, Rerimassie V, van Keulen I, Dorren G, Kaldenbach K (2014) Intimate technology : the battle for our body and behaviour. Rathenau Institute

Feher K, Katona AI (2021) Fifteen shadows of socio-cultural AI: a systematic review and future perspectives. Futures 132:102817

Ferrarello S (2019) The phenomenology of sex, love, and intimacy. Routledge. https://www.routledge.com/The-Phenomenology-of-Sex-Love-and-Intimacy-1st-Edition/Ferrarello/p/book/9780815358107

Gray FV (2016) Men's intrusion, women's embodiment: a critical analysis of street harassment (Routledge)

Hadreas P (2012) A phenomenology of love and hate. Ashgate Publishing

Hansmann U, Merk L, Nicklous MS, Stober T (2001) Pervasive Computing Handbook. Springer, Berlin Heidelberg
Husserl E (1973) Zur Phänomenologie der Intersubjektivität. Zweiter Teil vol 14. Martinus Nijhoff, pp 1921–1928
Ihde D (1990) Technology and the lifeworld. Indiana University, From garden to earth
Irwin S (2005) Technological other/ouasi other: reflection on lived experience. Hum Stud 28(4):453–467
Kanemitsu H (2019) The robot as other: a postphenomenological perspective. Philos Inquiries 7(1):51–61
Kudina O, Verbeek PP (2018) Ethics from within. Sci, Technol, Hum Values 016224391879371
Laato S, Hyrynsalmi S, Rauti S, Islam AKMN, Laine TH (2020) Location-based games as exergames—from pokémon to the wizarding world. Int J Serious Games 7(1):79–95
Leder D (1990) The Absent Body. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/A/bo3622735.html
Levy D (2007) Love and sex with robots: the evolution of human-robot relationships. Harper Collins
Liberati N (2015) Augmented "Ouch." How to create intersubjective augmented objects into which we can bump. In: 2015 IEEE international symposium on mixed and augmented reality—media, art, social science, humanities and design, pp 21–26
Liberati N (2017) Teledildonics and new ways of "being in touch": a phenomenological analysis of the use of haptic devices for intimate relations. Sci Eng Ethics 23(3):801–823
Liberati N (2018) Being Riajuu. A phenomenological analysis of sentimental relationships with "digital others." In: Love and sex with robots. LSR 2017 (Lecture notes in computer science). Springer, Cham, pp 12–25
Liberati N (2019a) Emotions and digital technologies. The effects digital technologies will have on our way of feeling emotions according to postphenomenology and mediation theory. Humana Mente 12(36):292–309. https://www.humanamente.eu/index.php/HM/article/view/297
Liberati N (2019b) Mediation theory between Pokémon GO and the everyday world. In: Geroimenko V (ed), Augmented reality games I understanding the Pokémon GO phenomenon, pp 51–60
Liberati N (2021a) Phenomenology and sex robots: a phenomenological analysis of sex robots, threesomes, and love relationships. Int J Tech 12(2)
Liberati N (2021b) La vita nell'oggetto in fenomenologia, postfenomenologia, nuovo materialismo e arte. Endoxa: Prospettive Sul Presente 29:97–100
Liberati N (2022) Digital intimacy in China and Japan: a phenomenological and postphenomenological perspective on love relationships at the time of digital technologies in China and Japan. Hum Studies Forthcoming 1–15
Liberati N, Nagataki S (2015) The AR glasses' "non-neutrality": their knock-on effects on the subject and on the giveness of the object. Ethics Inf Technol 17(2):125–137
Liberati N, Chen J (2022) Preliminaries to "digital intimacy" the exploration of the changes in values and meanings of intimacy in the digital age. BAU—Cont Contemp Cult 26:10–12. http://www.bauprogetto.net/quattordici.html
Liu J, Ouyang B (2022) A soul emerges when AI, AR, and anime converge: a case study on users of the new anime-stylized hologram social robot Hupo. New Media & Society, pp 1–23
Lu Z, Shen C, Li J (2021) More kawaii than a real-person live streamer: understanding how the otaku community engages with and perceives virtual youtubers. In: Conference on human factors in computing systems-proceedings, pp 1–14
Lunceford B (2010) Clothes make the person? performing gender through fashion. Commun Teacher 24(2):63–68
Pitsillides S, Jefferies J (2016) Intimate technologies: the ethics of simulated relationships situating ethics in technological futures. In: Proceedings of the 22nd international symposium on electronic. Art ISEA2016 Hong Kong, pp 144–151
Robertson R (1994) Globalisation or glocalisation? J Int Commun 1(1):33–52
Rosenberger R (2012) Embodied technology and the dangers of using the phone while driving. Phenomenol Cogn Sci 11(1):79–94

Sartre JP (2001) Being and nothingness : an essay in phenomenological ontology. Citadel Press
Sels L, Reis HT, Randall AK, Verhofstadt L (2021) Emotion dynamics in intimate relationships: the roles of interdependence and perceived partner responsiveness. In: Affect dynamics. Springer International Publishing, pp 155–179
Verbeek PP (2005) What things do. Philosophical reflections on technology, agency, and design. Penn State University Press
Verbeek PP (2011) Moralizing technology: understanding and designing the morality of things. University of Chicago Press
Viik T (2020) Falling in love with robots: a phenomenological study of experiencing technological alterities. Paladyn 11(1):52–65
Vosinakis S (2017) Digital characters in cultural heritage applications. Int J Comput Methods Heritage Sci 1(2):1–20
Wang J (2021) Should we develop empathy for social robots. In: Fan R, Cherry M (ed) Sex robots: social impact and the future of human relations. Springer, Dordrecht
Wellner G (2013) No longer a phone: the cellphone as an enabler of augmented reality. transfers 3(2):70–88. http://www.berghahnjournals.com/view/journals/transfers/3/2/trans030205.xml?pdfVersion=true
Yamaguchi H (2020) "Intimate relationship" with "virtual humans" and the "socialification" of familyship. Paladyn, J Behav Robot 11(1):357–369

Chapter 5
Augmenting Artificial Intelligences in Fiction: Evolving from Primordial Internet Memes to Cybergods of Disruption

Nathan Shafer

Abstract Beginning with a comparison between an augmented reality-based comic book series entitled *Wintermoot*, where Artificial Intelligence software is used in portions of the books and the first internet memes produced by the Meme-Rider Media Team (an art collective founded in 1999 in Anchorage, Alaska), this essay examines the usage of artificial intelligences in both the artistic sphere and in popular practice. The origin of the term 'meme' begins in Richard Dawkins's *The Selfish Gene* from 1976, where he describes the equivalent of genes in culture, the 'meme'. Dawkins ascribes several biological equivalents to memes from genetic theory, leading to the ability of memes to congregate together at a critical mass large enough to create an artificial intelligence. Case studies look at the way *Wintermoot* employs Artificial Intelligence in Augmented Reality throughout the comic book series. From here, a history of the usage of artificial intelligences as a cybernetic phenomenon in literature is compared with the way artificial intelligence are used in reality. Pulling from the early history of meme theory in nonfiction before the advent of the internet meme, this essay looks at the way memes propagate in society and how algorithmic thinking has produced social phenomenon beyond the sphere of the purely aesthetic and into a world of cybernetic networks. This is illustrated through the ways artificial intelligences are created (material negentropy), used and understood in literature and cinema. Comparisons are also made between the advent of social practice and mirror world theory in the art world, as well as the way memes evolved on the internet, concluding with a look at examples of artwork being produced in Augmented Reality that incorporates Artificial Intelligence as an artistic method for generating content, specifically the AIs in the *Wintermoot* series.

N. Shafer (✉)
Anchorage School District, Independent artist, Łuk'ae Tse' Taas Comics and Shared Universe, Structured Learning Classrooms, Alaska, USA
e-mail: admin@nshafer.com

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_5

5.1 Introduction

In the greater schema of artificial intelligence, there is a major riff between the speculative imagination and the physical manifestation. Artificial intelligences in the popular imagination anthropomorphize AI into characters that can be relatable or chaotic but are capable of reiterating and perpetuating new cultural ideas. AI in the physical world are a series of algorithms and programs that perform tasks based on the data that is fed to them. The AI in the physical world have learning algorithms built into them to help generate new and improved content but lack the ability to be culture bearers the way humans do.

Augmented reality functions in a very similar manner in the popular imagination, where the expectations of the technology are largely guided by the speculative musings in media. These expectations are not reflected in the physical manifestations of the technology. Augmented reality does not generate infinite amounts of content on demand, where everything in the world is a potential target for adventure and exploration. Artificial intelligence is not better at being human than humans are. Neither technology is going to undo the fabric of reality. How they function is completely contingent on the data we program them to work with.

There are literal, technical aspects of artificial intelligence in mixed reality environments that have specific functions. Consider augments that have an AI algorithm to obey the laws of physics. This is a useful version of the marriage of AI and AR, but the ones this chapter will look at are specifically in the arts and digital humanities. To do this, we have to return to a time when there were no internet memes, when memes were either a technical term a handful of zoologists in the world used, or an aesthetic concept used by artists.

5.2 The Meme-Rider Media Team

In 1999 Isaac Boatright, Joelle Howald and Nathan Shafer founded the Meme-Rider Media Team and began producing collaborative art projects derived from the basic notion of memes as defined by Richard Dawkins seminal work, *The Selfish Gene* (Dawkins 1976). These early meme-based artworks did not originally appear as what is now ‘internet memes’, but rather took the structure of interactive art, or what was a more prevalent description in the early 2000s, relational aesthetics, as proposed by Nicolas Bourriaud.

One of the first true ‘internet memes’ produced by the Meme-Rider Media Team (Fig. 5.1) is a reality hack of the Apple advertisement campaign, “Think Different”, where various ‘geniuses’ from history were pictured, with the caption, ‘Think different.’ Celebrity geniuses from the campaign included John Lennon, Albert Einstein, Amelia Earhart, Miles Davis, Pablo Picasso and several others. The meme produced by the Meme-Riders was of Hal 9000, the nefarious AI from Stanly Kupbrick’s *2001: A Space Odyssey*. This early internet meme was launched on the

Fig. 5.1 Hal 9000: Think different (Meme-Rider Media Team, 2007). Internet meme

seminal social media site MySpace, which would later give way to Facebook, Twitter and Instagram. When this internet meme was produced, augmented reality had not yet evolved into mobile augmented reality. It was dubbed several things, but was referred to as ‘locative media’, around 2005, the meme of Hal 9000 as an Apple advertisement for ‘Think Different’ was used as an internet link akin to ‘rick-rolling’.

Rick-Rolling was an internet phenomenon very prevalent on MySpace in 2005, where users would post links to various articles or sites, and the link would send other users to a video of a Rick Astley song, usually, ‘Never Gonna Give You Up’ or ‘Addicted to Love’. Rick-rolling and the ability to edit one’s own MySpace page with JavaScript to perform devious functions was one of the main reasons MySpace lost out to Facebook in the social media wars. Users were becoming frustrated with the perceived chaotic freedom inherent in social media, where fake accounts, fandoms and online tricksters had what felt like free range. Hard lines and algorithms in social media that put gates up between users paved the way for the manner in which social media is used at the writing of this essay (2022).

In 2009, when mobile augmented reality had become available to users via the Layar app, a version of the Hal 9000 meme was remade, where the image target of *Hal 9000: Think Different* would function as an augmented link to a Rick Astley song on YouTube. It functioned as an archive of the original spirit of the early internet meme.

The Meme-Rider Media Team produced several meme-based works during the 1999–2005 time period, while some of them were online, or internet based, many were not. The formats that internet memes evolved from were spread over several mediums and were defined at different stages across a spectrum of social philosophies. Dawkins original definition of meme, from *The Selfish Gene*, "a behavior, style or usage spread from person to person within a culture through imitation" (Dawkins 1976). Effectively the meme in memetics was the equivalent to the gene in genetics. The meme would be the genes of culture and would be affected by evolution and natural selection in a similar fashion.

The first really explicit usage of internet meme probably lies with Mike Godwin, when the internet was still in its infancy. As early as 1990 he was discussing the problems associated with online discussion groups which lead to his proposal which is now known as 'Godwin's Law'. Simply stated it infers that every online discussion group, as it grows, the probability of Naziism and comparisons to Hitler will occur. Meaning every online group, if it continues, Nazis and Hitler will come up in comparison. While this is not a certainty, Godwin's Law became a meme in the Dawkins tradition, spreading from person to person, evolving slightly, imperceptibly over time, until in 1993 Godwin proposed that internet memes, specifically his own meme, 'Godwin's Law', need to start being thought about more seriously as the internet community grows. Remember that this is 1993, one year before everyday users had reliable access to the internet through web-browsers.

5.3 AI/NI in Wintermoot, Material Negentropy

To unpack how we get from an early internet meme to various artificial intelligences existing in the reality spectrum, whether as works of art, or as transitional forms, we must first look at what we mean when we say, 'artificial intelligence'. The artificial intelligence this chapter will end with is called 'The Cybergod of Disruption'. It is a character in the augmented comic book series *Wintermoot* (Fig. 5.2). Inside the stories and worlds of *Wintermoot* there are a series of Artificial Intelligences and Natural Intelligences that more or less never really interact with the people the books are about. They are a kind of background radiation, always there, causing static, possibly inconsequential, always mysterious, both chaotic and uniform. The term 'natural intelligences' is not used in *Wintermoot*, but it is inferred. It is not antithetical or opposite of artificial/synthetic intelligences, but rather a function of biological/geological systems of the Earth. Manifestations similar to 'spirits' or supernatural entities connected to places, such as a 'god of the river' or 'spirit of the forest'. In essence they are more of a fantasy version of AI. More accurately they may be called *preternatural intelligences.*

In *Wintermoot*, the Natural Intelligences made themselves visible during a series of singularity events called Repocalypse, where an AI named 'Assistant to the Repocalypse' began rewriting reality using a series of five simultaneous weaponized singularities. The Natural Intelligences, Godwit Continuum and Wintermoot appeared and

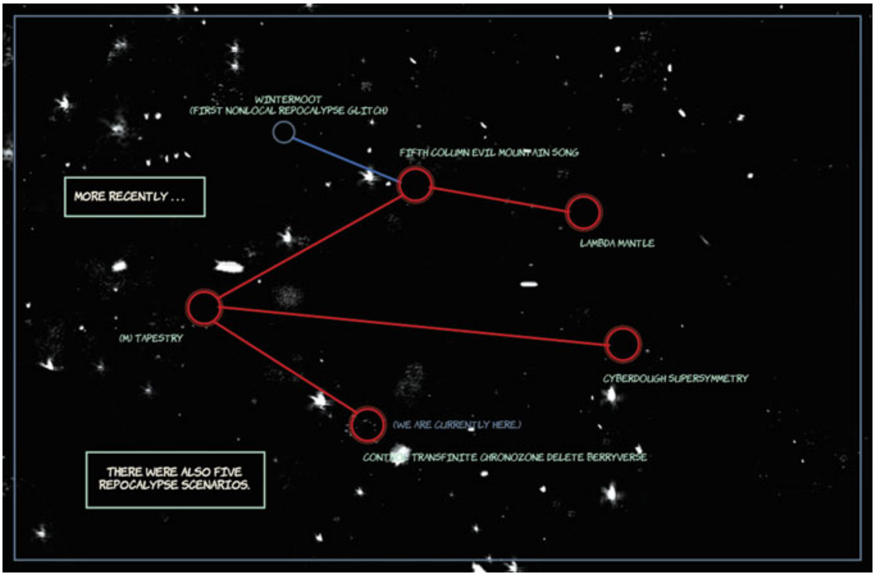


Fig. 5.2 Repocalypse panel from Wintermoot (Shafer 2019-2022). Augmented comic book

began restoring realities in the Multiverse, aided by several thousand superheroes from a smattering of alternative realities. The natural intelligences of *Wintermoot* are sentient manifestations of networks in nature, influenced heavily by cybernetics, information theory and James Lovelock's Gaia hypothesis.

While natural networks are entering the popular imagination more and more, they have about the same evidentiary footing as Gaia hypothesis and have been criticized by evolutionary biologists such as Stephen Jay Gould and the inventor of meme theory, Richard Dawkins. In terms of memes, however, the Gaia hypothesis is rife with its ability to mutate through natural selection in culture and turn into new thoughts. There is a mutualism, or cooperation, in Gaia hypothesis that many biologists actually concur with, but mostly out of hand, not based on any evidence. It does reflect however in a more Indigenized approach to the subject. Lovelock proposed that a mechanism on Earth could produce a kind of natural selection. Specifically, that nature would favor a species that does not damage the abiota of the planet. This notion runs parallel to an Indigenous way of thinking about the world, where humans have an obligation to creation, to steward it respectfully. Gaia proposes that the Earth favors species that upkeep, as stewards of itself. This has never been proven but is the kind of theory that functions better as a meme that can initiate other aspects of cultural evolution.

The natural network that is the most present in the popular imagination is the mycelium network, a subsurface network of fungi that interconnect with other organisms in a biome. It has even become a part of the Star Trek universe, with the spaceship in *Star Trek: Discovery* using the interstellar mycelium network to move the ship throughout the galaxy. For Lovelock, global mechanisms like the mycelial network, are in a neural network, or a brain. Howard Bloom in his book *Global Brain: Evolution of the Mass Mind from the Bing Bang to the 21st Century*, took this farther, pointing at several of the natural neural networks on the planet that function, or have functioned as 'brains'. (Bloom 2000) The brain is the thing here. It is what we mean when we talk about artificial intelligence. When we worry, love or feel indifferent to the notion that sometime in the future, robots will be able to think and feel as we do. That they will evolve a consciousness, emotional spectrum and the ability to do anything and everything humans can do, perfectly illustrated in the continuum of Human/Cylon beings in the *Battlestar Galactica* universe. Humans evolve, humans make robots, robots become humans and take over. It is, in essence, an existence paradox.

Returning to the global brains in *Wintermoot* for a moment, the Godwit Continuum is a network of biological movement connected to the Godwit bird which annually travels from northern Alaska to New Zealand. Wintermoot is the manifestation of all things human, gathering in the winter. The name Wintermoot is derived from multiple sources, the first being William Gibson's AI Wintermute from *Neuromancer*, the term 'moot' meaning both a 'gathering of people' and 'obsolescence' and finally the '-miut' suffix from Inuit languages meaning 'the people of'. In essence a 'wintermoot' in the comic book series is all of the who's, what's, when's, where's and how's of humans gathering in winter. In Alaska, where *Wintermoot* is set, this means storytelling and living off of the stores of food hunted and harvested before the snow falls. These are just some of the Natural Intelligences in the comic books. The intelligences represent natural cybernetics, sort of a naturalistic approach to cyberpunk aesthetics.

Antithetical to this are the Artificial Intelligences in *Wintermoot*, they are functions of material negentropy, the way humans organize nature and rewrite the laws of entropy, bringing anthropic order rather than the chaos which is the matter-of-course for the universe. Both of these intelligences are represented as shadows in *Wintermoot*, reflections of their true forms, which are imperceptible to humans. There are multiple ways these are illustrated, but they are designed to have the same type of shape that Dawkins' memes have, or what the internet really looks like. They are everywhere, nowhere in particular and can partially manifest in any ways that work. They are ordered interconnections, material negentropy, like food or houses. Pieces of the world brought to an order useful to humans, which is another philosophy akin to Indigenous Alaskan thinking, specifically in the Dena'ina worldview of southcentral Alaska.

In the Dena'ina worldview, at one point in the past, the animals could talk. They were shapeshifters and appeared different than they are today. According to the Dena'ina storyteller Peter Kalifornsky, Campfire People (the term for humans in the Dena'ina canon) could understand the animals and would interact with them this way. Eventually the animals decided to split off into pairs and make themselves useful to humans. As a story, this is when the animals entered into a state of material negentropy, they became ordered and useful. (Kalifornsky 1991).

To be sure, these are creative/speculative musings on both material negentropy and naturally occurring entropy, especially in relation to Dena'ina thinking. I am stating them here because science fiction tropes of sentient artificial intelligences have a rich history in literature, and the connection needs to be made between how we imagine AI to be and how it is. A connection should be kept too for the way the *Wintermoot* comics function, they are deeply connected to the worldview they were birthed in, southcentral Alaska, and Kalifornsky's retelling of the traditional Dena'ina story of "When the Animals Split into Pairs" is the basis for the way *Wintermoot* stories are presented, every book a different pairing of heroes meant to be useful to humans.

In the beginning, way back to when Ada Lovelace working on the Analytic Engine, especially to the work of Alan Touring on the Enigma Machines, computers were made to be useful to humans, a quintessential happenstance of material negentropy. When we measure a network's movement towards 'normalcy' we are measuring it with negentropy. Ultimately though, setting this binary of natural and artificial intelligences is a reflection of the binary of the way humans imagine the universe. It may be considered that these are not a binary, but a pairing, in the Dena'ina sense, that these seemingly different aspects are both useful to humans. Again, the notion of 'material negentropy' can illustrate this.

The technical notion of negentropy comes from Erwin Schrödinger's 1944 book *What is Lif*e? where he discusses the notion of negative entropy, where things move towards order, or normality; rather than the slow roll of entropy towards universal disorder. The natural world in our popular imagination is not used as a notion of entropy, but rather stereotypically aligned with the 'natural order'. Hence in the *Wintermoot* comic books, natural intelligences would be perceived as more positive and on the side of 'good' and the artificial intelligences, though they are ordering the world into a more useful normalcy for humans, are ultimately seen as chaotic and 'evil'.

Like computers themselves, AI was made to be useful to humans. At this moment in time AI are algorithms, programs, protocols. They are human-made systems and networks that can be programmed to learn and problem-solve, in essence all they really do is add IQ to any platform. They can also be employed to develop content, which is where we are currently operating. They are really nothing more, and simply augment human endeavors. AI software is designing images, text, music. They are producing content based on what a human artist is feeding it as data and the functions we are asking the software to perform. They have yet to truly produce memetic structures on their own as they have not been able to evolve a culture of their own, they are extensions, machinations, of our extended cultures.

Kevin Kelly noted the way AI was beginning to integrate with tech platforms "Amid all this activity, a picture of our AI future is coming into view, and it is not the HAL 9000—a discrete machine animated by a charismatic (yet potentially homicidal) humanlike consciousness—or a Singularitian rapture of superintelligence. The AI on the horizon looks more like Amazon Web Services—cheap, reliable, industrial-grade digital smartness running behind everything, and almost invisible except when it blinks off" (Kelly 2022).

5.4 The Book of Machines

Artificial intelligences have existed in literature and the popular imagination for long enough to have a history akin to evolution, albeit completely fictional. In fact, one of the earliest instances of an artificial intelligence in literature, or the popular imagination, was inspired by Charles Darwin's publication, *On the Origin of Species*, outlining evolution and speciation. The name of the book is *Erewhon, or Over the Range*, written by Samuel Butler in 1872. In the novel there are three chapters dedicated to the notion of machines evolving a consciousness, or at least a form of consciousness, obviously through natural selection. These chapters are famously known as *'The Book of Machines'*.

Even though this is one of the earliest notions of artificial intelligence, it is more accurately a form of synthetic consciousness which is an offshoot of artificial intelligence as an area of study, but it still plays into the underlying binary of the philosophical notions of vitalism and mechanism, which again function as a pairing just as well as a binary. Vitalism being more the notion that living things have a certain je nais se qua, while mechanism posits that natural networks are in essence just machines. Both vitalism and mechanism play into the fictionalizing of artificial intelligences as thinking/feeling robots whether nefarious or benign.

Vitalism here exhibits the same problematic teleological notions that Gaia offers. There is an unstated purpose, ghost in the machine, or intelligent design at play. While the mechanistic worldview tends to reduce everything in the universe to objects in a machine, which gives rise to notions such as ancient simulations, or the 'universe is a gigantic computer'. To be fair with these, both vitalism and mechanism in relation to fictionalizing of artificial intelligence, or incorporating AI into augmented reality art making, are completely viable, whether one agrees with them as philosophies or not, since the notions being produced are completely speculative. These issues were addressed further in *Wintermoot*, in the Cyberingian Simulation, an ancient simulation that sits as augmented reality on top of the physical world we are all experiencing. Cyberingia is neither in the past or future, but all at once, calculating and simulating all possibilities at the same time. Cyberingia itself though is several standard deviations for the artificial intelligence as it exists in the physical world we are currently experiencing.

In *Erewhon*, Butler is not talking about the artificial intelligences from today's popular imagination. If placed chronologically using literary punk, *Erewhon* is squarely rooted in the aesthetics of steampunk. In fact, the machines the writer is describing through these three chapters are 'vapour-engines', i.e., steam engines. There are philosophical musings on the engines being fed the way organisms must be fed, and that humans will ultimately end up being turned into slaves to appease said conscious vapour-engines. From Chapter Two in *Erewhon*, Butler states, "complex now, but how much simpler and more intelligibly organized may it not become in another hundred thousand years? or in twenty thousand? For man at present believes that his interest lies in that direction…here seem no limits to the results of accumulated improvements if they are allowed to descend with modification from generation to generation" (Butler 1872).

Butler is musing on the notion of machines reproducing themselves using humans the way flowers use bees. It is postulated that the machines will improve generation upon generation, and this is important here, as Butler quite correctly speculates that machines evolve opposite of natural organisms, in that they evolve to be more complex and smaller simultaneously, while the brains of individual organisms that exhibit the most consciousness evolve to be more complex and larger. What is interesting to note here are the notions of William Gibson's pairing of AIs from *Neuromancer*. They are siblings, *Neuromancer* housed in Rio de Janero and *Wintermute* housed in Bern. They are trying to merge to create a global brain, or superintelligence by gaining size, rather than diminishing. (Gibson 1984) Wintermute is also wanting to merge with another superintelligence outside of Earth. Achieving a world-wide network is a singularity event, which has been gone over multiple times, from Tielhard de Chardin's 'noosphere', to Kevin Kelly's 'hive mind' to Valentin Turchin's 'metasystem transition'. In the end, these notions really point to a fear of universalism in the human consciousness. We fear being drones in a colony, whether due to some deep evolutionary vestiges, like hypnagogic jerks while we are sleeping.

The final chapter of *The Book of Machines* has a return of the author from Erewhon, speaking on the machines there and the way they were perceived. The author notes that the machines can be seen as 'extra-corporeal limbs' of the humans who created them, and that ultimately man was a 'machinic mammal' in that they are the only species to extend their limbs beyond their bodies, they "lie about detached, now here and now there, in various parts of the world" (Butler 1872). Concluding however that there is no limb fashioned of any material that is better than a flesh and blood leg.

5.5 Mirror World, AI Winter

In the spirit of augmented reality, the reality of real-world artificial intelligences (AI in real life) is both more complicated and less fantastical. To explain, virtual objects are not beholden to the laws of physics. Early augmented reality would place virtual objects into the real world using either targets or geolocations. These objects were only visible via the screen of a smart device. These objects when viewed through the screen did not adhere to the ambient world of the device. They sat in front of all other objects or landscapes. They were not affected by gravity, light or heat. Their existence was contingent on a robotic eye, or lens to interpret the data that delineated their existence. In a manner of speaking, early augments were only tethered to reality via a screen. They are objects native to digital space. With the integration of artificial intelligence into mixed reality spaces, the real world started to become something that virtual/digital objects could now begin to interact with, albeit in a unilateral way.

Before giving examples of how these considerations materialize, artificial intelligence as a new media process should be placed in an appropriate context with the World Wide Web. The first World Wide Web launched a new era of 'global brain' reality. This first World Wide Web is where an interconnected set of websites, algorithms and protocols created a new reality of spaces across the globe and the human brains within it, similar to a 'superintelligence'. It was heavily influenced by the 'cyberspace' of *Neuromancer* and other elements of science fiction. Web2.0, also called the participatory web, was social media, the true birthplace of internet memes, avatars and a global 'social intelligence' that for better or worse created much of the reality the world is living with today, 'fake news' and 'clone accounts' included. Developers of the World Wide Web and Web2.0 became billionaires with these new spaces. For most elements of both of these webs, there is a long, rich history of speculative fiction that described it before it was a reality. For example, Neal Stephenson's *Snow Crash* from 1992 was using the terms 'avatar' and 'metaverse' well before social media and the World Wide Web.

Web3 is a bit different, but a place that is at once familiar and outside of the sphere of popular understanding. Web3 is a combination of mixed reality spaces and the blockchain. More importantly it is also emerging out of Stephenson's notion of the 'metaverse'. More accurately, Web3.0 is a materialization of David Gelertner's slightly earlier notion of the 'mirrorworld' from his book *Mirror Worlds: The Day Software Puts the Universe in a Shoebox…How It Will Happen and What It Will Mean*, (Gelertner 1992).

The notion of the 'mirrorworld' was once again popularized by futurist Kevin Kline at the eve of Web3, in 2019, noting that "the mirrorworld will appear to us as a high-resolution stratum of information overlaying the real world." (Kline 2019) The mirrorworld would be a virtual version of the real world, layered over top of the real world, via AR, VR, and any other variations thereof. But that is just the beginning, it is going to be bigger and more, a marriage of the virtual world, the physical world and all the things humans know, think and feel built into it. This space is where artificial intelligence will ultimately fit into the web platforms. The first two web platforms used algorithms to track and predict human behavior, systems that are built to learn by watching humans interact in the web platforms. We are constantly being mined for data.

AI, as of the writing of this chapter, is actually on the forefront of becoming viable. The speculative imagination that has defined AI is starting to get a real-world counterpart as artificial intelligence is actually being employed. During the first two web platforms, AI was not actually developing, our collective understanding of it was completely contingent on the fiction written about it. There was a lot of chatter about what it could do, coupled with very little invented, to move the technology forward, and even less in research funding to propel it. In 2014, as Web3 was on the far-off horizon, Kelly defined the three things in AI that were actually going to bring it into viability: Cheap parallel computation (GPUs from video games processing massive amounts of simultaneous virtual programs), Big data (the amount of data collected on the web becomes a data set) and better algorithms (being able to actually program and use neural nets) (Kelly 2022). This is all just to make the point that in reality and usage, AI is going to be pretty boring. There is not going to be an AI apocalypse. Artificial intelligence is going to be part of the ubiquitous computing universe.

What AI is going to be doing in these Web3 spaces is function as a reality fixer, or an agent of obedience, it will be part of the evolution of Web3, or maybe even the next platform. What AI will be able to do is mine all of the data around people, places, things and adjust the virtual object/ digital twin to obey the laws of physics as they are being experienced in the physical world. As has been discussed earlier in this text, virtual/digital objects do not obey the laws of physics. They simply exist in their own space, and when displayed via a networked device lay on top of the physical world. As augmented reality has evolved, sensors built into devices are able to 'read the room' and figure out where the ground is, various surfaces, even windows. They can place digital objects and make them obey the shape of curves of target objects that triggered them. Now consider digital objects that can respond to ambient temperature, the languages being spoken around the device, ambient smells (like in a cafe, or pizzeria) and even the specifics of the individual (user), things such as hunger, alertness or current thoughts.

Consider this in relation to the 'Internet of Things', the physical network of physical objects that are interconnected via internet enabled software. When every object can be a set of data that an AI can take, consider and re-present in the mirrorworld. The biggest goal would be to develop a way AI can interact with a virtual object to the point of having that virtual object reflect something into physical space. This might be a singularity point akin to the point where humans move from transhumans to posthumans, the next step in an evolutionary process, and this notion has been with mirrorworld since its inception.

From Gelertner's *Mirror Worlds*, "Consider Darwin's twin processes of speciation and evolution. Ensembles evolve; ensembles develop species. Individuals don't. …When you gather a bunch of such individuals together and allow them to interact, evolution and speciation emerge" (Gelertner 1992). Put this in relation to the notions of material negentropy, and there is an even greater sense of how humans order the world into useful things, then those things ignore reality (the laws of physics) then transcend them. The example that comes to mind here is the 5th dimensional architecture appearing in Christopher Nolan's film from 2014, *Interstellar*.

In the movie, Matthew McConaughey's character 'Cooper' has slipped into the event horizon of a black hole and is saved by a group of humans from the future who have created a multidimensional object that from the twenty-first century perspective would be ignoring the laws of physics, when in reality it is taking a virtual space and having it manifest in the physical world. In the movie this is illustrated when McConaughey/Cooper is able to visit the liminal space of a bookshelf in his daughter's bedroom, from there he virtually manipulates the books to send her a message in morse code.

5.6 Kauwarek, Cyberingia, Cybergod of Disruption

The Cyberingian Simulation of *Wintermoot* (Fig. 5.3) is part of the mirrorworld, albeit in an Indigenized fashion. The *Wintermoot Series* is part of a larger social practice project that collaborates in new media/speculative fiction spaces to initiate sustainable relationships that can create culturally appropriate comic books and graphic novels. This collaborative practice is being done by the Łuk'ae Tse' Taas Shared Universe. In creating the stories in this shared universe, artists are working to create speculative fiction that uses Indigenous Alaskan concepts.

One of the most important aspects of this thinking relates to the evolution of humans. The standard Occidental anthropological notion of human migration across the Earth states that humans evolved in Africa and migrated to the rest of the world. This scenario plays out in the Americas, with bands of people crossing from Asia via the Beringian Land Bridge. This is not the migration story told in Indigenous Alaska. There are several, in fact, and many of them have similarities, but notably all tell of very different movements and extrapolations. When the ocean levels were lower, during the Younger Dryas, people would have been living on the land bridge, not just walking from Asia into Alaska. There are stories of this place. In the Iñupiat tradition it is called *Kauwarek*, sometimes spelled 'Qawairaq'. The Iñupiat storyteller William A. Oquilluk wrote a version of these stories that he had heard called *The People of Kauwerak: Legends of the Northern Eskimo* in 1973. There are several elements in this collection that point at a very different world in Alaska than we see today, one where it was much warmer, there were different plants and trees growing there and

Fig. 5.3 Map of 'Cities' from Wintermoot (Shafer 2019-2022). Augmented comic book

the ancestors looked different. “Those long-ago people had a wide nose and mouth and bigger eyes than now. They had hairs on their bodies and long hair on their heads. Their bodies were weather-beaten and brown in color. They did not stay in one place. They had to rove around” (Oquilluk 1973).

Originally the world was so warm and abundant that the people had no wants, did not have to wear clothes and could live without worry. One day there was an earthquake and spirits emerged during the disaster, which is when the people started believing in them. One day it started getting colder and the snows started to come, eventually the people of Kauwarek invented whaling and fishing techniques as well as many other things to live in this new reality. The second disaster in this story is a great flood, which would chronologically coincide with the rising of the water, covering the Bering Land Bridge, putting the homelands of these ancient Iñupiat peoples underwater. The Kauwerak dialect of Iñupiatun is still spoken in Nome, Alaska—and there are stories that after the ‘Second Disaster’, the people spread out all along the shorelines from Alaska to the other side of Canada, and can still understand each other when they speak, but many new languages and dialects emerged.

This worldview is the basis from the way most of the Cyberingian Simulation of *Wintermoot* functions. It is a global reconstruction of every culture that the planet Earth has ever known. AIs in the story created a way of simulating all of these ancient peoples and societies and placing them in the real world, over top of physical reality. This simulation is in various layers and can give the appearance of traveling in time, but is really just accessing the simulation, or mirrorworld. Many of the cultures of Alaska also speak of giants who lived in Alaska before humans arrived. In the Dena’ina tradition, the ‘deghelay dnay’ (translates as ‘mountain people’), who were also shapeshifters, decided to become the animals, or mountains. The Dena’ina say that in the ancient times, it was a time when the animals could talk, they were shapeshifters, and the people could understand them. These realities are also part of the Cyberingian Simulation. There are aspects of the simulation that are completely out of time, the future, the present, the past, people can be half-human, shapeshift, move between worlds. And they can talk to the AIs that built this mirrorworld (Fig. 5.4a and b).

When designing the second book in the *Wintermoot Series*, Nathan Shafer and David Karabelnikoff were designing a group of cyberpunk Unangan characters that had built a futuristic metropolis/domed city in Adak called Sea Lion City. The city was filled with Indigenous AI designed using a Unangan-centric algorithm using a base-7 system (Ululung) and the GPT-2 AI software. The authors wanted to use an AI to create the dialogue the AIs would be using in the comic book. A system was devised to place as much of the printed Unangan worldview into an AI database with some selected texts on information theory, complex systems and ‘pataphysics. While the results did not produce the desired effects, it did stimulate the way text would be generated for both the Artificial and Natural Intelligences in the stories. The AI texts are not printed into the comic books and graphic novels, but rather are augmented over top of them (Fig. 5.5a and b).

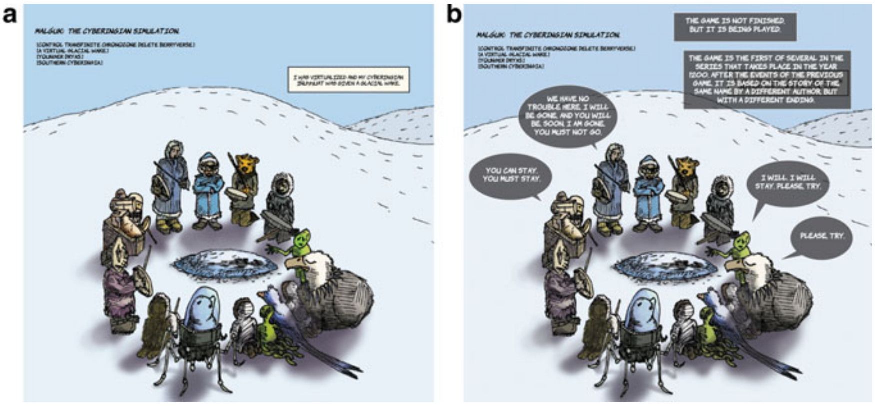


Fig. 5.4 (**a**) and (**b**) AI augments Wintermoot, Book One: Aqpik and Mars Apple (Shafer 2019-2022). Augmented comic book/AI text

The AI texts did not work for the graphic novel presentation of the story because what they generated was not able to take memes into consideration. The meme that was the biggest issue is the meme of 'cultural appropriateness'. The *Wintermoot* books and social practice of Łuk'ae Tse' Taas are purposefully created to exhibit best cultural practices and model culturally appropriate storytelling/cultural representation of Alaskan cultures. AI generated texts were not able to generate texts that did not read as culturally insensitive or at the very best, cultural appropriation. For example, one of the texts generated for the cyberpunk Unangan characters was the following, "That effect was found in men who had consumed alcohol at least three times a week for at least three years, but who were not breastfed." The AI is completely unaware of the history of alcohol in Tanam Unangaa (the Unangan homeland, also known as the Aleutian Chain). When the Russians began arriving in Tanam Unangaa, they would force men to hunt for sea otters by killing members of their family. They would also strand the hunters alone on islands, burning their iqyaqs (a kayak that was biometrically made for each individual hunter) leaving them with a barrel of vodka. The Russians and later the Americans had several waves of genocide committed against the Unangan people, in which alcohol was always weaponized. Casual dialogue discussing alcohol is inappropriate in this context.

Other texts from this process included dialogue such as, "We could feed us food that was not our own, that was fresh and fresh," or "The "Halo" is the name given to the Covenant vessel," which were used as augments.

Since the texts were not usable in printed graphic novels the way they were, they became connected with a chaotic force beyond the reality of the stories, an AI from the Cyberingian Simulation named the Cybergod of Disruption.

Every book from the *Wintermoot Series* has several versions of augments that sit over top of them. All of the images have a channel that will generate sets of AI images and dialogues over them, some of them from the Cybergod of Disruption, sometimes from something else. They are a part of the story to be sure but are also

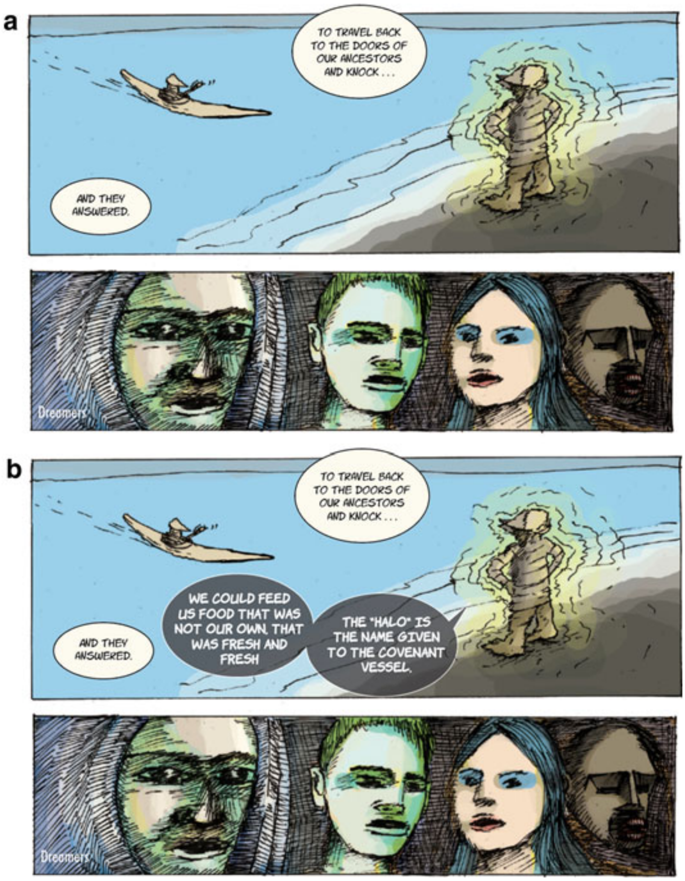


Fig. 5.5 (**a**) and (**b**) AI augments Wintermoot, Zero Chapter: Sngax-Six (Shafer 2019-2022). Augmented comic book/AI text

an aside that shows a hyper-fictionalized and AI generated set of stories that will begin to start manifesting themselves in the physical world of the comics the way that Cooper from *Interstellar* was able to interact with his daughter in the distant past. In essence, the Cybergod of Disruption's augmented AI-generated texts will begin to appear in the text as a sort-of fourth wall that refuses to obey the laws of the universe it is entering.

In *Wintermoot 4*, several characters begin meeting the AIs from the Cyberingian Simulation (Fig. 5.6a and b) The AIs appear as shadow kaiju beyond the horizon, not fitting into the reality around them, but still there. The AI can hear and see all of the other characters, but the others can only see them. The AI dialogue generated uses the GPT-2 platform to simply respond to the dialogue that is being read in the comic books (Fig. 5.7a and b).

a

b

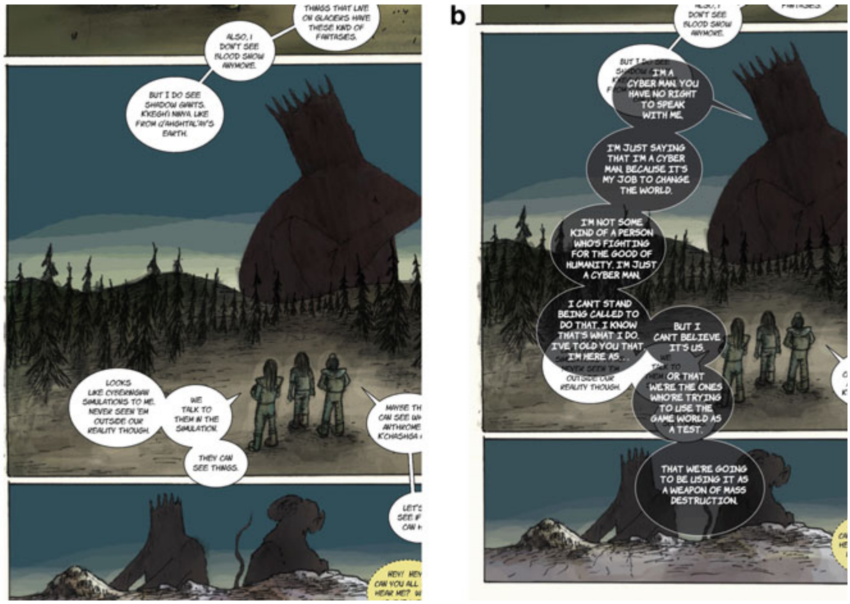


Fig. 5.6 (**a**) and (**b**) AI augments Wintermoot, Book Four: Anthrome and K'chashga (Shafer 2019-2022). Augmented comic book/AI text

5.7 Conclusion

As AI merges more into the mirrorworld of Web3, we will see more of what the actual, usable functions are. Consider not thinking of the physical manifestation of the real world and its mirrorworld equivalent as binaries or opposites, but rather in the Dena'ina sense of pairing, specifically that they are pairings useful to humans, the way we have constructed material negentrooy of the world around us. Ultimately it is a very human endeavor and extension of ourselves into the worlds around us.

Since the writing of this paper, AI has had an explosion of users online, specifically focused around the Lensa mobile app, which uses the Stable Diffusion AI software. There are ethical concerns being raised about the way Stable Diffusion and hence Lensa is using the databases that run their programs. Namely, that there are over 5 billion images that Stable Diffusion has extracted from the internet and used in their software to create visual art. The author is not condoning or promoting AI image generators that are using ethically questionable methods, but rather focusing on how AI can be seen within the reality spectrums we are all living in. AI is going to be a vital part of Web3 and what it evolves into next, as we continue to our evolution with material negentropy. It is not always going to be fair and thoughtful, as the internet has shown, going all the way back to the beginning of this paper where

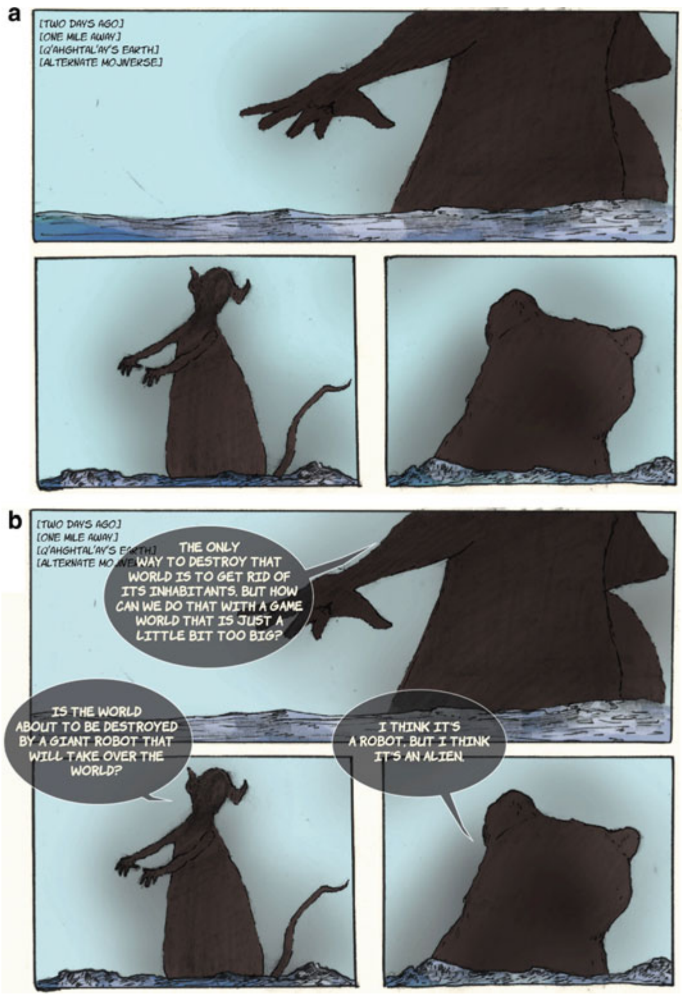


Fig. 5.7 (**a**) and (**b**) AI augments Wintermoot, Book Four: Anthrome and K'chashga (Shafer 2019-2022). Augmented comic book/AI text

Rick-Rolling was discussed. Rick Astley has never received royalty payments from the billions of Rick-Rolls performed online since the early 2000s, Like the Cybergod of Disruption in *Wintermoot*, AIs can function as both fictional/theoretical entities while simultaneously running AI software in real time, as mirror world equivalences.

References

Bloom H (2000) Global brain: evolution of the mass mind from the Bing Bang to the 21st Century. John Wiley & Sons, New York

Butler S (1872) Erewhon, or over the range. https://www.gutenberg.org/files/1906/1906-h/1906-h.htm. Accessed 15 Oct 2022
Dawkins R (1976) The selfish gene. Oxford University Press, New York
Gelertner D (1992) Mirror Worlds: the day software puts the Universe in a shoebox…How It Will Happen and What It Will Mean. Oxford University Press, New York
Gibson W (1984) Neuromancer. Ace Science Fiction Books, New York
Godwin M. Meme, counter-meme. https://www.wired.com/1994/10/godwin-if-2/. Accessed 15 Oct 2022
Kalifornsky P (1991) A dena'ina legacy k'tl'egh'l sukdu: the collected writings of peter kalifornsky. Alaska Native Language Center, Fairbanks
Kelly K (2022) AR will spark the next big tech platform–call It Mirrorworld. https://www.wired.com/story/mirrorworld-ar-next-big-tech-platform/. Accessed 15 Oct 2022
Oquilluk W (1973) The people of Kauwerak: legends of the Northern Eskimo. http://www.alaskool.org/native_ed/historicdocs/people_of_kauwerak/kauwerak_pp.htm. Accessed 15 Oct 2022
Schrödinger E (1944) What is Life?
Shafer N (2019–2022) Wintermoot. Łuk'ae Tse' Taas, Anchorage

Part II
The Educational Use of Intelligent Augmented Environments

Chapter 6
Artificial Intelligence, Augmented Reality and Education

Alba Rusillo-Magdaleno, Alberto Ruiz-Ariza, Sara Suárez-Manzano, and Teresa Martínez-Redecillas

Abstract This chapter proposes several tools based on Artificial Intelligence (AI) and Augmented Reality (AR) to promote the transmission of knowledge in the educational context. AI is already used to provide assistance as part of educational processes (Squirrel AI, Google Classroom) and can generate effective methodological and assessment resources (Alexa, SIRI, NVIDIA Canvas, WriteToLearn, Lingvist, Duolingo, Gradescope, Smartick). In addition, the combination of AI and AR facilitates the learning of scientific, geographical, mathematical and social content through the use of interactive scenarios (Metaverse, Munzee, Goosechase, Blippar, PlugXr or ActionBound, Bodyplanet, Arloon Geometry, Chromville Science, Explore the World). The objectives of this chapter are to analyse the functionality, advantages and limitations of the use AI and AR in an educational environment, to illustrate its practical applicability and to create a table/guide with didactic examples.

6.1 Introduction

Technological advances represent the main characteristic of society in the twenty-first century (Johnson and Wetmore 2021). At the forefront of these new technologies is artificial intelligence (AI), which is capable of making predictions and automated decisions based on the processing of large amounts of data (Shrestha

A. Rusillo-Magdaleno · A. Ruiz-Ariza (✉) · S. Suárez-Manzano · T. Martínez-Redecillas
Group HUM-943: Physical Activity Applied to Education and Health, Faculty of Humanities and Educational Sciences, University of Jaen, Jaen, Spain
e-mail:arariza@ujaen.es

A. Rusillo-Magdaleno
e-mail: arusillo@ujaen.es

S. Suárez-Manzano
e-mail: ssuarez@ujaen.es

T. Martínez-Redecillas
e-mail: tmartinezredecillas@gmail.com

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_6

et al. 2019). In addition, augmented reality (AR) can create a new digital space that opens the way for three-dimensional interactions between the real and the virtual (Xiong et al. 2021).

Both AI and AR currently have global influence, and this is especially observable not only in the fields of management and communication, but also in the everyday life and education of young people (Iatsyshyn et al. 2020), which has led to the emergence of a new educational paradigm that takes into account the demands of a global, interconnected and multicultural society (Calero 2019). This situation implies the need for teachers to undergo continuous training in new technologies, and to adapt to a new model of communication with students (Vlasova et al. 2019). Current teaching approaches need to be transformed to take advantage of these new digital tools, and new approaches to planning, organisation and assessment need to be created to compensate for the possible limitations of these systems (Colás et al. 2018).

The use of information and communication technologies is now accepted as a way to devise novel, creative and participatory teaching strategies for students; if used appropriately, they can be a very efficient teaching tool (Martzoukou et al. 2020). As a consequence, the role of the teacher changes, and becomes that of a mediator between learners rather than the exclusive bearer of knowledge. This opens up the possibility of more dynamic and interactive teaching, which allows children and adolescents to construct their own knowledge by investigating, exploring and playing (Thomes 2019).

The most recent studies indicate that programmes based on the use of these software and digital devices can motivate and satisfy students and improve the teaching–learning process, regardless of the subject matter (Cabero-Almenara and Puentes-Puente 2020; Irhebhude et al. 2019). In the following, some of the main educational tools based on AI and AR, the advantages and limitations of these in terms of didactic use, and a brief review of the practical applications for students are presented.

6.2 Artificial Intelligence

AI is the ability of a machine to make decisions based on previously learned or acquired data, thus simulating the abilities of humans (Rouhiainen 2018). When this technology is used, it must meet at least three requirements; it must be (i) lawful (i.e., obliged to comply with laws and regulations), (ii) ethical (i.e. committed to respecting the ethical principles of human autonomy), and (iii) robust (i.e. considered a solid tool in terms of applicability) (Smuha 2019).

AI has proven to be useful in all fields of knowledge. Its functions include recognising, classifying and labelling objects in images; improving algorithmic business strategy; processing patient data; predictive enterprise maintenance; distributing content on social networks; protecting against threats to internet security; and performing tasks that would be complex or dangerous for humans (Rouhiainen 2018).

However, all these forms of usability have undergone major transitions within the various domains. In education, early computer designs have evolved into integrated systems (robots, smart classrooms) that work in parallel with the teacher to assist students in the teaching–learning process (Chen et al. 2020). These applications have facilitated distance education, improved administrative and institutional performance, and allowed for more complex and efficient feedback processes based on the immediate analysis of large amounts of data provided by students. In addition, AI has enabled the development of automated programming, improved knowledge systems and pattern recognition, and machine learning and language processing, among other achievements (Zhang and Lu 2021). However, AI does not only involve logical thinking, and future developments in this field are already moving towards new forms of emotion-based AI. This would be the first step towards a machine intelligence that would surpass the intellectual capacity of humans.

Although AI has unquestionably advanced the well-being of modern societies, there are associated privacy issues that may affect the personal information of users of this technology. These arise due to the massive exposure of personal data such as spoken language, identity, biographical data or location on online platforms (Akgun and Greenhow 2021). These issues require specific legislation and rules with the aim of protecting such data and preserving the security of individuals.

6.2.1 Educational Applications, Advantages and Limitations of Artificial Intelligence

The main functions of AI, in terms of their impact on student learning, are associated with the virtual decision making they offer through relevant data obtained from an educational environment (Selwyn et al. 2022). This technology has great teaching potential, as it is capable of speeding up task completion processes, reducing difficulties in accessing learning, automating management processes, and improving learning outcomes (Moreno-Padilla 2019).

The advantages resulting from the use of AI are diverse. For example, AI is able to analyse and track the educational progress of students, and to identify gaps among them in terms of learning (Chen et al. 2022; Seo et al. 2021). The use of AI also makes it possible to select human talent (Reyes et al. 2021; Tippins et al. 2021), detect learning deficits (Hosseini et al. 2022), make decisions automatically, and drive personalised and collaborative learning (Rahimi et al. 2022). These advantages arise because AI generates intrinsic motivation for learning, while at the same time increasing active classroom participation and student interaction (Chan et al. 2022). In addition, many of these educational applications allow for assessment through intelligent multiple-choice tests. AI creates response patterns, and is able to verify learning based on an initial diagnosis and predicted prognosis, all of which efficiently reinforce the criterion-based assessment that should predominate with children (Moreno-Padilla 2019).

However, in order to understand the educational dimensions of AI, it is first necessary to determine what skills-based AI is. This is defined as AI that facilitates collaboration and provides assistance to teaching–learning processes and can be used to develop methodological and evaluation tools. According to Lara-García (2022), Benbya et al. (2020) and Teigens et al. (2020), skills-based AI can be classified using the following typology:

- Artificial narrow intelligence: This a tool that can strengthen a person's skills; for example, AI systems can process and execute tasks and large amounts of data, and can allow people to handle situations that require emotion management, creativity, intuition or empathy.
- General artificial intelligence: This is characterised by the ability to learn or understand intellectual tasks that can be performed by humans.
- Strong artificial intelligence or super intelligence: The aim in this case is to equal human intelligence. Although this is still a long way from being achieved, researchers aspire to experiment with consciousness.

The use of this technology is not without certain drawbacks, and these must be taken into account. Some of the most important limitations include the scarcity of resources in schools, a lack of teacher training, and dependence on these technologies by students. Moreover, since these drawbacks directly affect the people who use them, training and protection systems must be provided that ensure human dignity, as well as physical and mental integrity (Sánchez-Bravo 2020). Employers using AI media should ensure that their workers do not suffer bias, discrimination or stigmatisation. Finally, as a technology, AI should never lead to the deception of a user or limit his or her ability to make choices; the user must have the ability at all times to object to the choices made by the AI (Sánchez-Bravo 2020). Table 6.1 summarises the specific advantages and limitations of a selected range of AI-based educational technology tools.

6.2.2 Technological Tools for the Use of AI in the Educational Context

Practical educational applications based on AI include intelligent conversational software (better known as chatbots), online platforms for self-learning, and educational robotics (Moreno-Padilla 2019). Chatbots are computer programs that interact with users through conversational interfaces (Zierau et al. 2021). Online self-learning platforms make it possible to direct the interaction of teachers and learners towards the achievement of educational goals, according to the personal needs of each user (Liu 2022). Finally, educational robotics is mainly oriented towards automatic task design (Ionescu and Schlund 2021). Its information complements new teaching methods very efficiently, and favours more active learning by students (Vivas-Fernández and Sáez-López 2019). In this way, AI contributes to the educational community through

Table 6.1 AI-based tools for educational use

Squirrel IA http://squirrelai.com/	This is a virtual platform that offers extracurricular classes and personalised tutoring. It can reinforce a student's learning through online courses or assistance within their school. Recommended age: 6–12 years (Fig. 6.1)
Advantages	Ease of use Identification of learning gaps Allows for personalisation of content
Limitations	Does not allow students to intervene in their own teaching–learning process Low privacy of minor users due to data exposure
Example of practical application	**Subject**: language
	Aim(s): reading, taking into account various aspects (intonation, rhythm, speed and accuracy), understanding written texts, applying reading comprehension strategies, and developing critical skills
	Activity: students read a written text, after which they complete a questionnaire on the concepts covered. Based on the student's answers, Squirrel provides the teacher with an evaluation, which summarises the content that needs to be reinforced
Google Classroom https://classroom.google.com/u/0/	This is a virtual platform that incorporates exercises, interactive lessons and comments, allowing students to put into practice the knowledge they have acquired in an autonomous and personalised way. This application allows teachers to remotely monitor the teaching–learning process. Recommended age: > 10 years (Fig. 6.2)
Advantages	Ease of use Free-to-use platform Centralised storage of exercises, lessons, assignments and comments in Google Drive
Limitations	Lack of teacher guidance in activities Lack of automated diagnostic tests
Example of practical application	**Subject**: mathematics
	Aim(s): solving problems using procedures and strategies that require planning and application of mathematical operations
	Activity: students download a list of problems that must be solved using the acquired solving strategies. The student needs to upload the results to Google Classroom within a specific time, so that the teacher can proceed to the evaluation

(continued)

Table 6.1 (continued)

Alexa https://developer.amazon.com/es-ES/alexa	This is a voice-controlled virtual assistant that includes spoken commands. Some of the services offered by Alexa are audiobook reading, a calculator, a translator, a virtual encyclopaedia and mathematics practice. These functions involve dynamic learning of multiple linguistic concepts. Recommended age: > 6 years (Fig. 6.3)
Advantages	Ease of use 24-h support, every day of the week Skills can be installed that enhance the functionality of the device
Limitations	Absence of parental or teacher control Incompatibility with some devices
Example of practical application	**Subject**: natural sciences
	Aim(s): recognising certain ecosystems and the living beings that inhabit them
	Activity: students engage in conversation with the Alexa device. The app briefly and simply describes some of the ecosystems. The student needs to recognise the system and give the answer in the shortest possible time. When the student has answered all the questions related to the topic in question, Alexa provides a reward in the form of a song or story
SIRI https://www.apple.com/es/siri/	This is a voice-controlled personal assistant. SIRI is integrated into a mobile device and is capable of helping, organising, complementing or supporting students in the activities they have to carry out. It offers personalised solutions based on the dynamics of learning progression. Recommended age: > 10 years (Fig. 6.4)
Advantages	Ease of use Instant resolution of doubts Reduced teaching load
Limitations	Absence of cognitive empathy Absence of privacy, due to continuous functioning and continuously recorded speech
Example of practical application	**Subject**: French
	Aim(s): identifying general information from short oral texts. The topics can be familiar and drawn from everyday contexts related to one's own experience

(continued)

Table 6.1 (continued)

	Activity: students prepare and provide SIRI with questions in French on a given topic. This technological tool answers the questions and poses additional questions to the learner. The child then asks a question or describes an object, and Siri has to guess what it is. Subsequently, the roles are reversed
Nvidia Canvas https://www.nvidia.com/es-es/studio/canvas/	This is a graphics processor for the gaming market. It can provide real educational experiences by bringing power and performance to a student's work. Nvidia Canvas is a tool that creates different realistic landscapes based on small sketches. Recommended age: > 10 years (Fig. 6.5)
Advantages	Ease of use Creative process enhancer Short editing times
Limitations	Difficult to install Requires a high-priced NVIDIA Geforce RTX graphics card
Example of practical application	**Subject**: climate science
	Aim(s): identifying the characteristics of atmospheric weather, weather stations and characteristics of the climate
	Activity: working in pairs, one of the students describes the weather or a climate of their choice through mime. The other partner uses the application to guess what is being described. The roles are then exchanged. The winner is the student who has best characterised their drawing in terms of what they have described
WriteToLearn	This is a diagnostic website that identifies and determines a student's weaknesses and needs in the areas of reading and writing. It allows the teacher to help each student to improve their writing, on an individual basis. Recommended age: > 8 years (Fig. 6.6)
Advantages	Ease of use Provides motivation for writing skills Provides comprehensive assessment and feedback based on the progress of each individual learner
Limitations	Website available only in English, Spanish and Chinese Web application not free of charge
Example of practical application	**Subject**: language

(continued)

Table 6.1 (continued)

	Aim(s): planning and writing clean, clear text by following certain steps, including drafting, revising and improving
	Activity: the teacher acts out a story through mime. The pupils are asked to write their interpretation of what the teacher is presenting, in a clean, clear and orderly way. The tool gives feedback with instructive comments about the writing and the story
Lingvist	This is a language learning platform that analyses students' skills and knowledge in real time. It uses large databases to adapt the contents to the level of the users. Recommended age: 6–16 years (Fig. 6.7)
Advantages	Ease of use Use of advanced language learning techniques Offers the opportunity to learn and review new words
Limitations	Premium payment platform Compatible only with certain mobile phones
Example of practical application	**Subject**: English
	Aim(s): identify, understand and use the linguistic structures and lexical aspects of a foreign language
	Activity: through a quiz, students complete cards with shapes, pictures or words. Students associate the content of the cards with the vocabulary they have previously acquired. Lingvist provides information on the language learning progress in the form of a percentage
Duolingo	This is a web platform for autonomous language learning. Using gamification, it analyses students' learning patterns and offers support in weaker areas. Recommended age: 4–16 years (Fig. 6.8)
Advantages	Ease of use Offers dynamic and fun learning Provides the possibility of studying anywhere and at any time
Limitations	Lack of a teacher to answer questions online Absence of elements focusing on grammar, text comprehension and phonetics
Example of practical application	**Subject**: English
	Aim(s): participating in different communicative situations and interacting on everyday topics, including vocabulary learned from different contexts

(continued)

Table 6.1 (continued)

	Activity: students complete vocabulary and grammar questions presented via the platform, and can achieve different levels. The tool then makes a learning diagnosis and determines the students' progress
Gradescope	This is a web platform that allows teachers to carry out diagnostics and evaluation processes remotely, while guaranteeing high-quality evaluations. Recommended age: > 18 years (Fig. 6.9)
Advantages	Configuration is based on preset criteria related to efficiency, objectivity, equity and transparency Provides statistics or rubrics for the evaluation process
Limitations	Complexity of use Requires a high level of prior training of the teacher in the use of the platform Interface only available in English and Spanish
Example of practical application	**Subject**: natural sciences
	Aim(s): identifying and interpreting knowledge related to the main parts of a machine, based on its structure and operation
	Activity: students take a multiple choice test on the parts of a machine and how they work. The teacher is responsible for semi-automatic evaluation of the documents
Smartick	This is an application aimed at mathematics learning. The website adapts the exercises shown to the students in real time, based on the detected level of ability. Recommended age: 4–14 years (Fig. 6.10)
Advantages	Ease of use Promotes learning with positive reinforcement in the form of games and pastimes Daily notifications to parents or tutors of sessions not completed by students
Limitations	Excessively short sessions (15 min) Paid application only
Example of practical application	**Subject**: mathematics
	Aim(s): identifying incorrect results in basic mathematical operations
	Activity: students complete various mathematical operations or solve problems at the different levels presented by the application. Evaluation information is displayed in real time to allow the student to modify incorrect answers

its impacts on the cognitive development of students in the areas of research, critical thinking, decision making, and problem solving, as well as through the communication and teamwork that result from its use (Alimisis 2013). Table 6.1 and Figs. 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9 and 6.10 summarises several AI tools that were designed as methodological resources and have promoted the development of effective learning among students. In other cases, the tools described here can suggest efficient methods of assessment by teachers.

Fig. 6.1 Use of the Squirrel Platform. *Source* https://www.youtube.com/watch?v=tmxKBsBL_6M

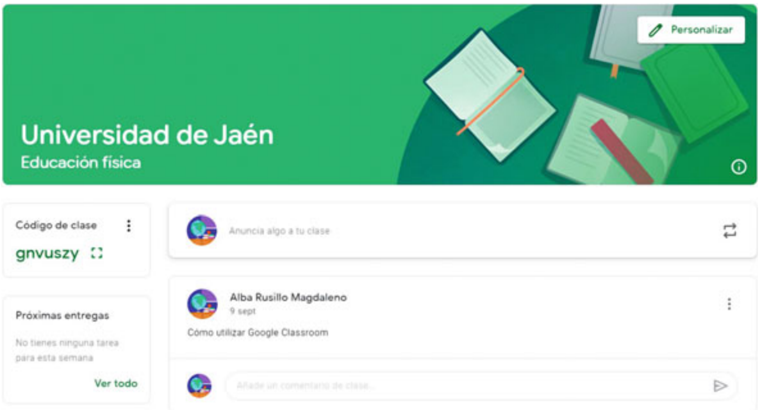


Fig. 6.2 Use of Google Classroom. *Source* Authors' own screenshot

Fig. 6.3 Use of Alexa. *Source* https://youtu.be/OqrW-BnUuME

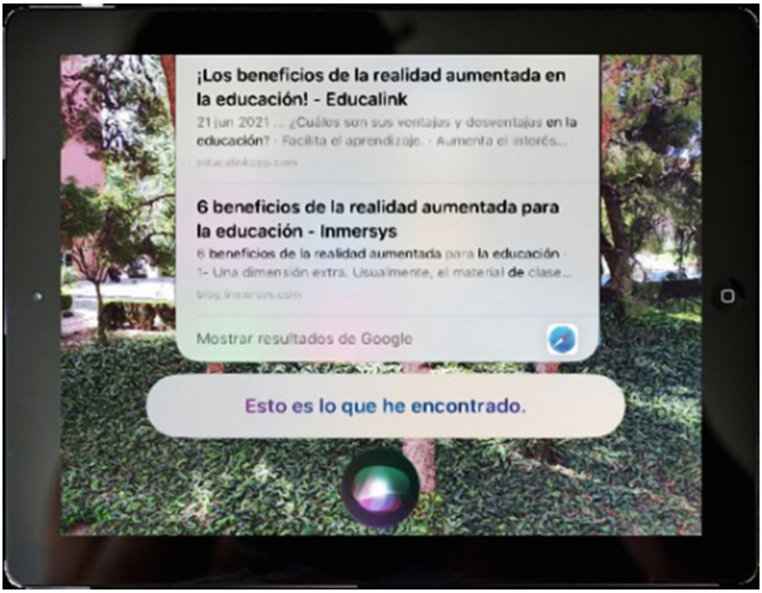


Fig. 6.4 Use of Siri. *Source* Authors' own screenshot

6.3 Augmented Reality

AR involves the creation of an artificial technological environment that combines real components with virtual elements (Gurevych et al. 2021). A projection is created via software installed on a device, which allows for the enrichment of reality through interaction with virtual elements (Cabero et al. 2018). For correct and effective use of AR, the device must be composed of various elements, such as a camera (to perceive information coming from the real world), hardware (which combines reality with

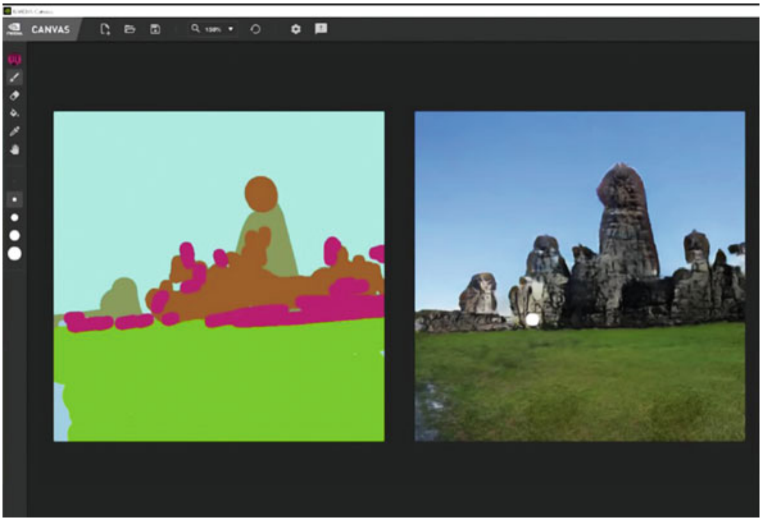


Fig. 6.5 Use of Nvidia Canvas. *Source* https://www.youtube.com/watch?v=mlZYRwJ2oJg

Fig. 6.6 Use of WriteToLearn. *Source* https://www.youtube.com/watch?v=KNxHuIgCpI0

virtuality), software (to manage the coupling process), and a marker (which allows the elements generated in the production to be projected) (Hamzah et al. 2021).

In general, AR can be classified into three types (Cabero et al. 2018; Viglialoro et al. 2021):

(i) physical AR, which may include several levels, such as (a) black and white patterns, (b) images or illustrations, (c) environments and (d) points determined by coordinates;

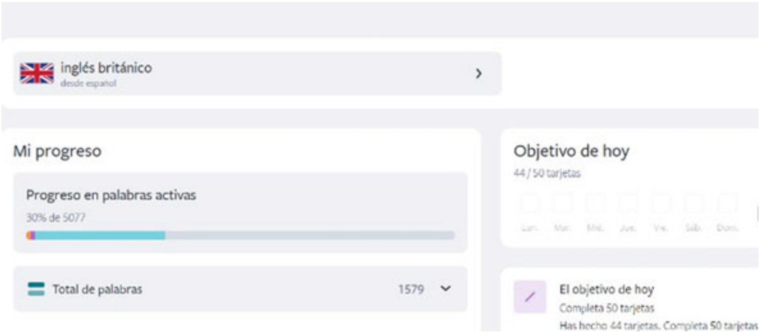


Fig. 6.7 Use of the Lingvist platform. *Source* Authors' own screenshot

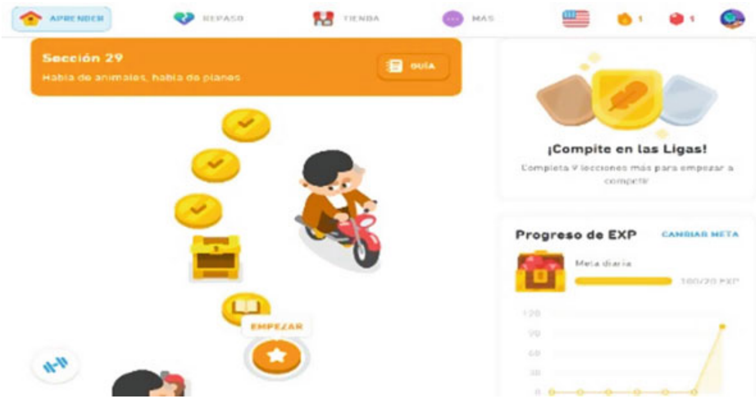


Fig. 6.8 Use of Duolingo. *Source* Authors' own screenshot

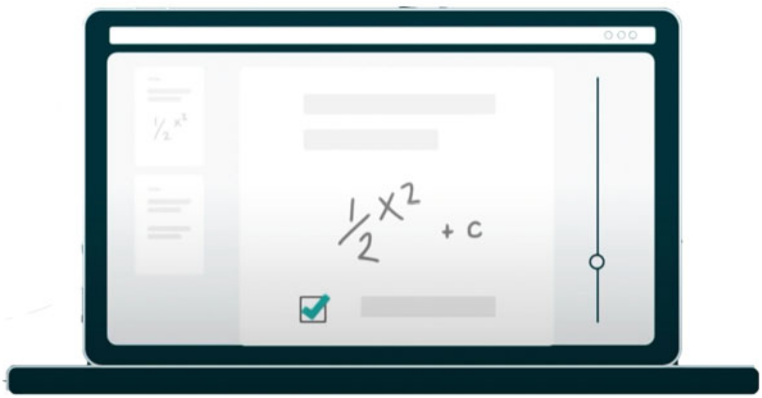


Fig. 6.9 Use of Gradescope. *Source* Authors' own screenshot

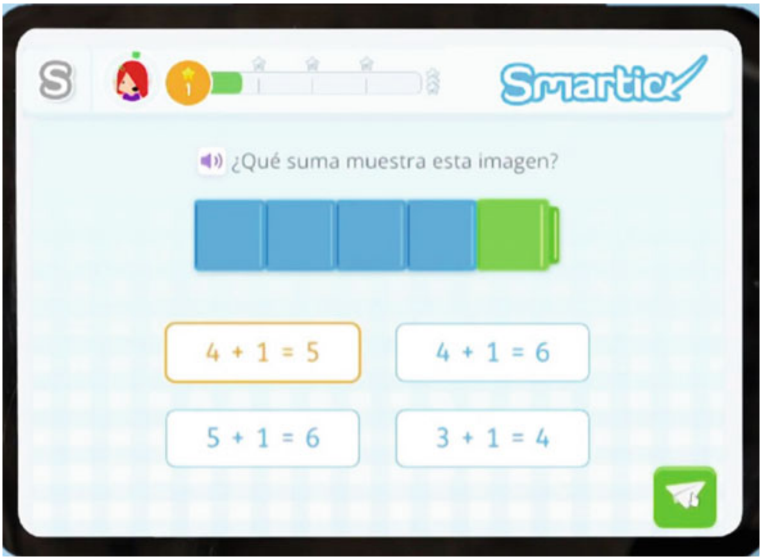


Fig. 6.10 Use of Smartick. *Source* https://youtu.be/739eQ5CAP5w

(ii) virtual AR, which involves mixed reality, where the dimensions of video, text, images and audio are presented at the same time;
(iii) functional AR, which allows for correct knowledge of the context to make the right decisions, create environments or artificial environments.

As a technology that increases interactivity with what is to be communicated (Kumar 2021), AR has a wide range of functionalities. This technology can complement the messages displayed by the media with graphics, videos or QR codes (Wang et al. 2021). In the domain of sports, AR can recognise variables related to participants or resources, thus allowing training sessions to be adapted (Le Noury et al. 2022). In medicine, AR can complement the training of medical students with virtual patients, by overlaying information about procedures (Yeung et al. 2021). In the context of industry, AR can provide tools for simulating the assembly and disassembly of a machine in a real environment (Baroroh et al. 2021). In the field of tourism, AR can be integrated with a location tracking function for the creation of interactive travel guides to provide information on tourist destinations and monuments (Wu and Lai 2021), and in the automotive industry, AR allows designers to improve their structures and to make comparisons through virtual representations (Chen et al. 2019).

Despite these advantages, however, AR poses certain risks due to the misuse of users' privacy. Some applications collect biometric data without consent, through facial analysis, which can lead to identity theft (Cowan et al. 2021). Personal data, phone status or location, contact lists, and recordings of all kinds can be stored and used fraudulently (Harborth et al. 2019). For example, some common applications access phone resources when the applications themselves are not running (Harborth et al. 2019). Users are therefore strongly recommended to review the privacy policies of AR apps, especially in the case of minors, as well as the ethical standards that developers must meet regarding physical, psychological, moral and social aspects (Steele et al. 2020).

6.3.1 Educational Applications, Benefits and Limitations of Augmented Reality

Among the many applications of AR is its use in education. This technology involves virtual simulators that allow for the creation, presentation and manipulation of objects from different perspectives, which can motivate learners (Osadchyi et al. 2021). More specifically, AR-based applications have shown a strong capacity to support learning through interaction with other elements of the curriculum. Over the last decade, the inclusion of AR in the classroom has begun to be consolidated, and is associated with highly innovative teaching methods due to greater enrichment of information, a better focus of learning towards what is considered substantial, improvements in attention (Tang et al. 2021), and a more appropriate contextualisation of content (Cabero et al. 2018). Consequently, these technological tools can promote positive educational performance, and have proven to be effective in the development of attention, concentration, memory, reasoning, motivation and interest (Marín-Díaz 2018).

The benefits of AR in the classroom are manifold. Programmes using AR have been shown to increase motivation for learning by improving the spatial perception of the objects of study (Mystakidis et al. 2021). When used in secondary education, AR facilitates the understanding of complex phenomena, enables interaction with virtual objects through the manipulation of real objects, and improves the development of autonomous and collaborative learning (Rivadulla and Rodríguez 2020). AR has also been found to enhance inclusive education and to offer significant benefits to students with learning difficulties (Kaimara et al. 2021).

However, although AR is considered a useful, scalable tool that can assist in teaching functions (Chan et al. 2022), it has some limitations. These include the difficulty of use (Fombona and Vázquez-Cano 2017), privacy and data protection issues, the high cost of the devices, the requirements for teacher training, and its low adaptability for students with educational needs (Kaimara et al. 2021). Table 6.2 summarises the specific advantages and limitations of some AR-based educational technology tools.

Table 6.2 AR-based tools for educational use

Metaverse	This is a mobile application that allows for the creation of school learning activities. Each user can create an avatar and interact with other users. Recommended age: > 9 years (Fig. 6.11)
Advantages	Ease of use Allows for interpersonal relationships between participants Provides an intuitive learning environment
Limitations	Privacy, security and user protection issues, due to the fact that crimes in this virtual ecosystem cannot be regulated Application is only available in English
Example of practical application	**Subject**: languages
	Aim(s): analysing the lexicon of one's own language, classifying words according to their category, and progressively improving in using them to express common situations
	Activity: Students are asked to collect words by getting a specific number. Each task proposed by the teacher provides a key to obtaining a new word
Munzee	This is a mobile application that allows the user to create geolocation games. Munzee provides students with QR codes that can be used to complete a given activity. Recommended age: > 8 years (Fig. 6.12)
Advantages	User-friendly application Application involves physical movement
Limitations	Requires the use of very large spaces Minimum spacing of 45 m between QR codes
Example of practical application	**Subject**: social sciences
	Aim(s): learning to interpret simple maps and plans to reflect interpersonal locations
	Activity: the students, working in groups, travel through a given space interpreting maps and directions provided by QR codes to collect all the elements of the solar system. Each code provides instructions that will make the participants analyse them and to orient themselves in space to reach another place, obtain another clue and collect another astronomical element. The winner is the group that has oriented itself best and collected all the elements that make up the solar system
Goosechase	This is a mobile application that combines orientation in space with smartphone technologies. The application allows the user to create searches through geolocation in a given context. Recommended age: > 8 years (Fig. 6.13)
Advantages	Creation of personalised experiences Promotion of interpersonal relationships among students
Limitations	Requires one game licence per event Confusing experience setup procedure

(continued)

Table 6.2 (continued)

Example of practical application	**Subject**: mathematics
	Aim(s): learning the meaning of decimal numbers through operations and problem solving
	Activity: numerical Gymkhana: students solve numerical missions with decimals found within a space. Each test provides a series of digits to be used in the final activity, which involves combining all the numerical digits from the previous activities to obtain a code. This acts as a key to open a lock and to get a reward chosen by the teacher in charge of the activity
Blippar	This is a mobile application used to create experiences by adding markers and multimedia elements. The tool generates interactive and playful content. Recommended age: > 4 years (Fig. 6.14)
Advantages	Creation of experiences without the need to download a software package Ease of use, management and creation of experiences
Limitations	Available only for iPhone and iPad Downloading this application takes up a lot of space on the device
Example of practical application	**Subject**: natural sciences
	Aim(s): investigating and learning about the anatomy of the human body and how it functions
	Activity: the application shows students the organs of the human body. Some of these organs are shown in detail, and students are required to answer a multiple choice question; other organs are hidden by shadows, and the students are asked to guess what they are. To see the complete anatomy and how it works, the students solve puzzles related to the topic being worked on, to obtain clues and unlock the hidden parts
PlugXr	This is a cloud-based platform that uses remote servers connected to the internet to store, manage and process data. The functionality of this tool consists of creating and projecting immersive augmented reality experiences, without the need for coding concepts. Recommended age: > 18 years (Fig. 6.15)
Advantages	Ease of use Creation of personalised experiences by the teacher
Limitations	Paid platform Available only for Iphone and Ipad
Example of practical application	**Subject**: English
	Aim(s): understanding messages in a foreign language, using the information conveyed to carry out specific tasks
	Activity: the classroom becomes a small city, in which clues are displayed in English in the form of concrete words or objects. When all the clues are found, students complete an interactive game in which pictures and words appear and need to be matched according to their meaning

(continued)

Table 6.2 (continued)

Actionbound	This is an application designed to create interactive tours that reinforce the content that is being worked on in the classroom. Recommended age: > 4 years (Fig. 6.16)
Advantages	User-friendly application Allows the teacher to analyse the results of the students' activities in real time
Limitations	Paid platform for schools Little personalisation and characterisation in the subject matter of the tests
Example of practical application	**Subject**: mathematics
	Aim(s): identifying and classifying representations of geometric figures
	Activity: ScapeRoom: this game is made up of different stations. At each of these stations, a QR code is scanned that shows different puzzles, activities, riddles, or geometric figures. As the activities are completed by the students, the teacher provides a QR code fragment. To get the final reward, students combine the QR fragments in order to open the door to the room
Body Planet	This is a mobile application that focuses on learning about the anatomy and functioning of the human body. The application allows the user to create digital environments using real objects. Recommended age: > 4 years (Fig. 6.17)
Advantages	Useful and highly motivating experiential learning experience Stronger understanding of abstract concepts
Limitations	Licence valid on only one device Additional equipment (T-shirts) is required
Example of practical application	**Subject**: natural sciences
	Aim(s): identifying, classifying and explaining the functioning of the organs, apparatus and systems of the human body
	Activity: the class is divided into groups, each of which has a T-shirt and a device with the Body Planet app. One member of each group puts on the T-shirt, and another focuses on the body to observe what comes out of the situation. When the students observe the organ, apparatus or system that appears on the screen, they classify it and give an oral explanation of its functionality
Arloon Geometry	This is a mobile application that allows the user to study geometry by observing figures from different perspectives. AR can be used to manipulate the polyhedra to improve spatial vision. The app presents multiple choice questions to evaluate what has been learned. Recommended age: > 4 years (Fig. 6.18)
Advantages	Highly visual, meaningful and interactive learning Effectively combines conceptual and procedural knowledge
Limitations	Paid application Does not show feedback after wrong answers from students

(continued)

Table 6.2 (continued)

Example of practical application	**Subject**: mathematics
	Aim(s): identifying and describing geometric bodies
	Activity: each student is asked to represent a different geometric body on a piece of cardboard. Using the Arloon Geometry app, students observe the geometric body in three dimensions and complete a card describing its characteristics. Finally, all the cards are hung on the wall and the bodies are associated with their characteristics
Chromville Science	This application supports science learning by combining creativity, motivation and technology to accelerate the educational process of students. Recommended age: > 4 years (Fig. 6.19)
Advantages	App respects the privacy of users Immersive experience that provides a very enriching experience for students
Limitations	Precise use of supplementary payment material Loss of colour in the representations
Example of practical application	**Subject**: natural sciences
	Aim(s): identifying and relating the organs, apparatus and systems involved in the vital functions of the human body
	Activity: students fill in and colour in a card related to the contents of the human body and its functions. Using the Chromville app, they observe the human body (skin, muscles and bones); finally, they choose one of these categories and build a model that represents what they have selected
Explore the World	This app shows monuments from all over the world in the form of three-dimensional models. Students are required to answer questions correctly at all levels to become an explorer. Recommended age: 9–11 years (Fig. 6.20)
Advantages	Free app download Available for IOS and Android
Limitations	Requires paid supplementary material Synchronisation problems with devices
Example of practical application	**Subject**: social sciences
	Aim(s): learning about and identifying the artistic works and historical monuments of Europe in a fun and playful way
	Activity: in groups, students study a particular country. Using the app, they explore the world and the complementary globe. Students identify, note down and then describe the most important historical monuments found in the territory assigned to them

6.3.2 *Technological Tools for the Use of AR in the Context of Education*

As can be seen from the above discussion, AR is considered beneficial for learning and practice, as its multimedia content is particularly motivating for learners (Khan et al. 2019). In the field of education, AR has been designed to respond to multiple learning demands, such as increasing the level of physical activity (Ruiz-Ariza et al. 2018) and improving the quality of writing (Wang 2017), mathematical competence (Sommerauer and Müller 2014) or foreign language learning (Hsu 2017). The efficiency of this type of learning is evidenced by significant improvements in collaborative learning (Wen 2021) and knowledge retention over the medium to long term (Alzahrani 2020). The use of these devices has also been found to increase student engagement (Alzahrani 2020) and learning satisfaction (Xue et al. 2019).

As noted above, the application of AR in education involves the creation and delivery of academic content through the use of new technologies to complement traditional learning methodologies. In a very general way, tools based on AR can be classified into development tools and reading tools (Osadchyi et al. 2021). The formers are used to create AR applications (Chen et al. 2019), while the latter allow users to interact with 3D models of hard-to-imagine objects and processes (Kiryakova 2020). Practical AR-based applications are most commonly delivered via handheld electronic devices, such as smartphones or tablets, although desktop devices can also be used to create or configure AR experiences. New designs for educational AR tools are also moving towards the creation of robots that are capable of presenting 3D content in addition to diagnosing problems, planning and assessing learning (Diehl et al. 2020). In all cases, current designs can be adapted to any environment, including vehicles, schools and homes (Udayan et al. 2020; Wiegand et al. 2019). Table 6.2 and Figs. 6.11, 6.12, 6.13, 6.14, 6.15, 6.16, 6.17, 6.18, 6.19 and 6.20 summarises several tools, based on the use of AR as a methodological resource, that have demonstrated the development of effective learning among students. In other cases, the tools presented here suggest efficient methods of assessment for teachers.

6.4 Conclusion

The domains of application of recent technological advances are numerous and have become particularly prominent in the first quarter of the twenty-first century. The current uses of AI and AR include those oriented towards the education of young people. These technological tools can improve the teaching–learning process, motivate students and foster interpersonal relationships. Today, parents and education professionals can choose from a wide range of tools based on AI and AR to complement the learning of their children or pupils. Although these tools have proven to be effective in a multitude of areas, their features and educational benefits are still little known; in general, there is still a belief in the field of education that this is

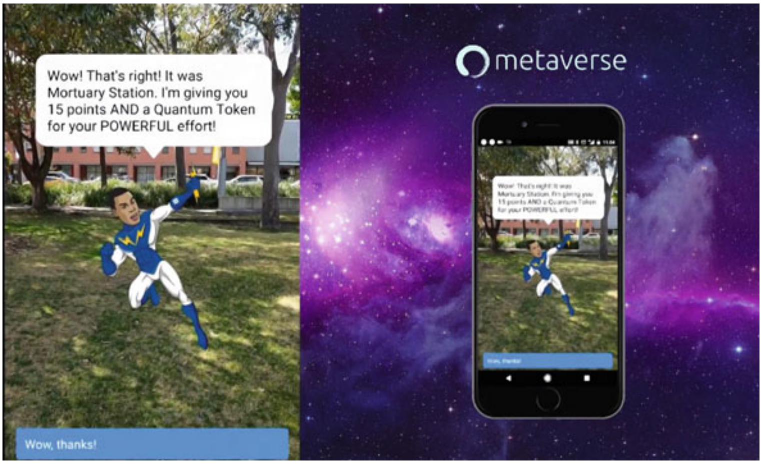


Fig. 6.11 Use of the Metaverse app. *Source* https://youtu.be/q_xPNpRCyLc

Fig. 6.12 Use of the Munzee app. *Source* https://www.youtube.com/watch?v=6yGZTojS5OM

an expensive, complicated and inaccessible form of technology. This chapter has highlighted some evidence from studies supporting the didactic potential of AI and AR in the teaching–learning process of young people. Our review of a wide range of didactic tools based on AI and AR may assist in an appropriate selection of these technologies. The reader can compare the main advantages and disadvantages of each application and its suitability in terms of the ages and abilities of the learners. Finally, although these are novel applications and much research remains to be done, no evidence of adverse effects from their use in the field of education has been found.

Fig. 6.13 Mission created with Goosechase. *Source* https://youtu.be/jAo7vND9z9I

Fig. 6.14 Experience of AR with Blippar. *Source* https://www.youtube.com/watch?v=-CikXblyGFg

Fig. 6.15 Creating AR experience with PlugXr. *Source* https://www.youtube.com/watch?v=yIanc3t3PyA

Fig. 6.16 Experience of AR with ActionBound. *Source* https://www.youtube.com/watch?v=FXj3e-krNDQ

Fig. 6.17 Experience of AR with Body Planet. *Source* https://youtu.be/QMLsXi99udE

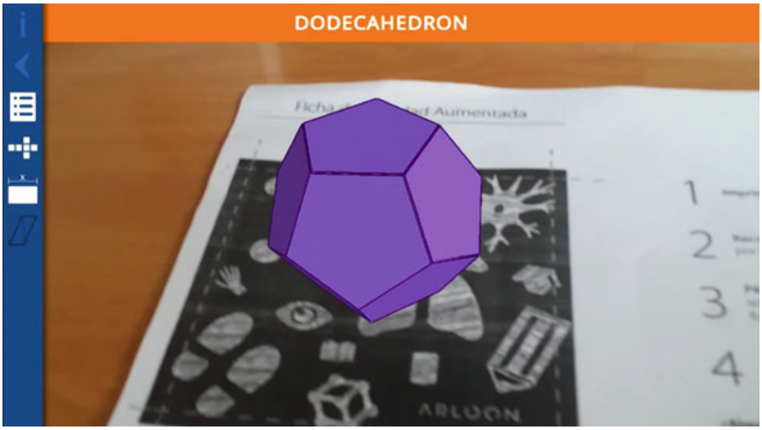


Fig. 6.18 AR experience with Arloon Geometry. *Source* https://youtu.be/F7gHsloqZM4

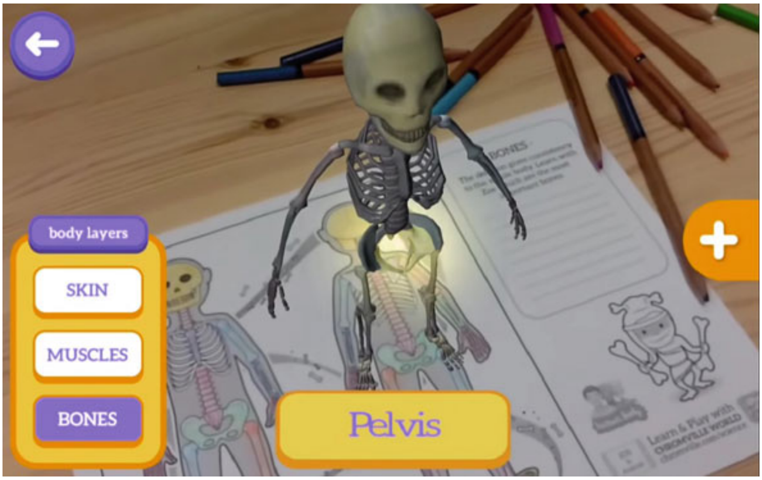


Fig. 6.19 Experience of AR with Chromville Science. *Source* https://youtu.be/qhykI82UiMc

Fig. 6.20 Experience of AR with Explore the World. *Source* Authors' own screenshot

References

Akgun S, Greenhow C (2021) Artificial intelligence in education: addressing ethical challenges in K-12 settings. AI Ethics 2(1):1–10. https://doi.org/10.1007/s43681-021-00096-7

Alzahrani NM (2020) Augmented reality: a systematic review of its benefits and challenges in e-learning contexts. Appl Sci 10(16):5660

Alimisis D (2013). Educational robotics: open questions and new challenges. Themes Sci Technol Educ 6(1):63–71. http://earthlab.uoi.gr/theste

Baroroh DK, Chu CH, Wang L (2021) Systematic literature review on augmented reality in smart manufacturing: collaboration between human and computational intelligence. J Manuf Syst 61(1):696–711. https://doi.org/10.1016/j.jmsy.2020.10.017

Benbya H, Davenport TH, Pachidi S (2020) Artificial intelligence in organizations: current state and future opportunities. MIS Q Executive 19(4)

Cabero-Almenara J, Puentes-Puente A (2020) La Realidad Aumentada: tecnología emergente para la sociedad del aprendizaje. AULA, Revista de Humanidades y Ciencias Sociales 66(2):35–51

Cabero J, de la Horra I, Sánchez J (2018) La realidad aumentada como herramienta educativa. Ediciones Paraninfo, SA

Calero CCS (2019) La llegada de las nuevas tecnologías a la educación y sus implicaciones. Int J New Educ 4(1). https://doi.org/10.24310/IJNE2.2.2019.7449

Chan L, Hogaboam L, Cao R (2022) Artificial intelligence in education. Applied artificial intelligence in business. Springer, Cham, pp 265–278

Chen L, Chen P, Lin Z (2020) Artificial intelligence in education: a review. IEEE Access 8(1):75264–75278

Chen X, Zou D, Xie H, Cheng G, Liu C (2022) Two decades of artificial intelligence in education. Educ Technol Soc 25(1):28–47

Chen Y, Wang Q, Chen H, Song X, Tang H, Tian M (2019) An overview of augmented reality technology. J Phys: Conf Ser 1237(2):022082

Colás MP, de Pablos J, Ballesta J (2018) Incidencia de las TIC en la enseñanza en el sistema educativo español: una revisión de la investigación. Revista de Educación a Distancia (RED) 18(56). https://doi.org/10.6018/red/56/2

Cowan K, Javornik A, Jiang P (2021) Privacy concerns when using augmented reality face filters? Explaining why and when use avoidance occurs. Psychol Mark 38(10):1799–1813. https://doi.org/10.1002/mar.21576

Diehl M, Plopski A, Kato H, Ramirez-Amaro K (2020) Augmented reality interface to verify robot learning. In: 29th IEEE international conference on robot and human interactive communication (RO-MAN), pp 378–383

Fombona J, Vázquez-Cano E (2017) Posibilidades de utilización de la geolocalización y realidad aumentada en el ámbito educativo. Educación XXI 20(2):319–342. https://doi.org/10.5944/educXX1.10852

Gurevych R, Silveistr A, Mokliuk M, Shaposhnikova I, Gordiichuk G, Saiapina S (2021) Using augmented reality technology in higher education institutions. Postmod Open 12(2):109–132

Hamzah ML, Rizal F, Simatupang W (2021) Development of augmented reality application for learning computer network device. Int J Interact Mob Technol 15(12)

Harborth D, Hatamian M, Tesfay WB, Rannenberg K (2019) A two-pillar approach to analyze the privacy policies and resource access behaviors of mobile augmented reality applications. In: Proceedings of the 52nd Hawaii international conference on system sciences. University of Hawai'i Press, Honolulu, HI, pp 5029–5038

Hosseini M, Saugstad M, Miranda F, Sevtsuk A, Silva CT, Froehlich JE (2022) Towards global-scale crowd + AI techniques to map and assess sidewalks for people with disabilities. arXiv preprint arXiv:2206.13677

Hsu TC (2017) Learning English with augmented reality: do learning styles matter? Comput Educ 106(1):137–149. https://doi.org/10.1016/j.compedu.2016.12.007

Iatsyshyn AV, Kovach VO, Lyubchak VO, Zuban YO, Piven AG, Sokolyuk OM, Iatsyshyn AV, Popov OO, Artemchuk VO, Shyshkina MP (2020) Application of augmented reality technologies for education projects preparation. In: 7th workshop on cloud technologies in education. http://ceur-ws.org/Vol-2643/paper07.pdf

Ionescu TB, Schlund S (2021) Programming cobots by voice: a human-centered, web-based approach. Procedia CIRP 97(1):123–129

Irhebhude ME, Abdullahi F, Agwi UC, Kolawole AO, Mohammed L, Tech B (2019) Augmented reality: user perception on effectiveness in educational delivery assessment in Kaduna. Pac J Sci Technol 20(2):102–111

Johnson DG, Wetmore JM (2021) Technology and society: building our sociotechnical future. MIT Press

Kaimara P, Deliyannis I, Oikonomou A, Fokides E, Miliotis G (2021) An innovative transmedia-based game development method for inclusive education. Digit Cult Educ 13(2)

Khan T, Johnston K, Ophoff J (2019) The impact of an augmented reality application on learning motivation of students. Adv Hum-Comput Interact 2019(1). https://doi.org/10.1155/2019/7208494

Kiryakova G (2020) The immersive power of augmented reality. Human 4.0—from biology to cybernetic. IntechOpen

Kumar TS (2021) Study of retail applications with virtual and augmented reality technologies. J Innov Image Process (JIIP) 3(2):144–156

Lara-García J (2022) Inteligencia Artificial y Justicia. DIVULGARE Boletín Científico De La Escuela Superior De Actopan 9(17):41–46. https://doi.org/10.29057/esa.v9i17.8093

Le Noury P, Polman R, Maloney M, Gorman A (2022) A narrative review of the current state of extended reality technology and how it can be utilised in sport. Sports Med 1–17

Liu Y (2022) Study on the blended teaching mode of analog electronic technology based on chaoxing learning platform. Adv Vocat Tech Educ 4(1):49–54

Marín-Díaz V (2018) Augmented reality in the service of educational inclusion. Case study. Retos XXI 2(1):60–72. https://doi.org/10.33412/retoxxi.v2.1.2060

Martzoukou K, Fulton C, Kostagiolas P, Lavranos C (2020) A study of higher education students' self-perceived digital competences for learning and everyday life online participation. J Doc 76(6):1413–1458

Moreno-Padilla RD (2019) La llegada de la inteligencia artificial a la educación. Revista de Investigación en Tecnologías de la Información: RITI 7(14):260–270. https://doi.org/10.36825/RITI.07.14.022

Mystakidis S, Christopoulos A, Pellas N (2021) A systematic mapping review of augmented reality applications to support STEM learning in higher education. Educ Inf Technol 1(1):1–45

Osadchyi VV, Valko NV, Kuzmich LV (2021) Using augmented reality technologies for STEM education organization. J Phys: Conf Ser 1840(1)

Rahimi SA, Cwintal M, Huang Y, Ghadiri P, Grad R, Poenaru D, Gore G, Zomahoun HTV, Légaré F, Pluye P (2022) Application of artificial intelligence in shared decision making: scoping review. JMIR Med Inform 10(8):e36199

Reyes MD, Gómez A, Ramos EV (2021) Desafíos de la gestión del talento humano en tiempos de pandemia COVID 19. Revista Universidad y Sociedad 13(6):232–236

Rivadulla JC, Rodríguez M (2020) La incorporación de la realidad aumentada en las clases de ciencias. Contextos educativos: revista de educación 25(1):237–255. https://doi.org/10.18172/con.3865

Rouhiainen L (2018) Inteligencia artificial. Alienta Editorial, Madrid

Ruiz-Ariza A, Casuso RA, Suarez-Manzano S, Martínez-López EJ (2018) Effect of augmented reality game Pokémon GO on cognitive performance and emotional intelligence in adolescent young. Comput Educ 116(1):49–63

Sánchez-Bravo ÁA (2020) Marco Europeo para una inteligencia artificial basada en las personas: European framework for people-based artificial intelligence. Int J Digit Law 1(1):65–77. https://doi.org/10.47975/IJDL/1bravo

Selwyn N, Rivera-Vargas P, Passeron E, Puigcercos RM (2022) ¿Por qué no todo es (ni debe ser) digital? Interrogantes para pensar sobre digitalización, datificación e inteligencia artificial en educación. SocArXiv. https://doi.org/10.31235/osf.io/vx4zr

Seo K, Tang J, Roll I, Fels S, Yoon D (2021) The impact of artificial intelligence on learner–instructor interaction in online learning. Int J Educ Technol High Educ 18(1):1–23
Shrestha YR, Ben-Menahem SM, Von Krogh G (2019) Organizational decision-making structures in the age of artificial intelligence. Calif Manag Rev 61(4):66–83
Smuha NA (2019) The EU approach to ethics guidelines for trustworthy artificial intelligence. Comput Law Rev Int 20(4):97–106
Sommerauer P, Müller O (2014) Augmented reality in informal learning environments: a field experiment in a mathematics exhibition. Comput Educ 79(1):59–68. https://doi.org/10.1016/j.compedu.2014.07.013
Steele P, Burleigh C, Kroposki M, Magabo M, Bailey L (2020) Ethical considerations in designing virtual and augmented reality products—virtual and augmented reality design with students in mind: designers' perceptions. J Educ Technol Syst 49(2):219–238
Tang YM, Chau KY, Kwok APK, Zhu T, Ma X (2021) A systematic review of immersive technology applications for medical practice and education-trends, application areas, recipients, teaching contents, evaluation methods, and performance. Educ Res Rev 100429
Teigens V, Skalfist P, Mikelsten D (2020) Inteligencia artificial: la cuarta revolución industrial. Cambridge Stanford Books
Thomes M (2019) El uso de las nuevas tecnologías en el ámbito educativo. In: XI Congreso Internacional de Investigación y Práctica Profesional en Psicología. XXVI Jornadas de Investigación. XV Encuentro de Investigadores en Psicología del MERCOSUR. I Encuentro de Investigación de Terapia Ocupacional. I Encuentro de Musicoterapia. Facultad de Psicología—Universidad de Buenos Aires, Buenos Aires
Tippins NT, Oswald FL, McPhail SM (2021) Scientific, legal, and ethical concerns about AI-based personnel selection tools: a call to action. Pers Assess Decis 7(2):1
Udayan JD, Kataria G, Yadav R, Kothari S (2020) Augmented reality in brand building and marketing—valves industry. In: 2020 international conference on emerging trends in information technology and engineering (ic-ETITE), pp 1–6
Viglialoro RM, Condino S, Turini G, Carbone M, Ferrari V, Gesi M (2021) Augmented reality, mixed reality, and hybrid approach in healthcare simulation: a systematic review. Appl Sci 11(5):2338
Vivas-Fernández L, Sáez-López JM (2019) Integración de la robótica educativa en Educación Primaria. Revista Latinoamericana de Tecnología Educativa. RELATEC 8(1):107–128. https://doi.org/10.17398/1695-288X.18.1.107
Vlasova EZ, Avsentieva EY, Goncharova S (2019) Artificial intelligence—the space for the new possibilities to train teachers. Artif Intell 40(9):17
Wang TH, Kao CH, Wang TJ (2021) Implementation of mobile learning in mathematics instruction for elementary second graders. Mathematics 9(14):1603
Wang YH (2017) Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. Comput Educ 113(1):162e176. https://doi.org/10.1016/j.compedu.2017.04.013
Wen Y (2021) Augmented reality enhanced cognitive engagement: designing classroom-based collaborative learning activities for young language learners. Educ Technol Res Dev 69(2):843–860
Wiegand G, Mai C, Holländer K, Hussmann H (2019) Incarar: a design space towards 3d augmented reality applications in vehicles. In: 11th international conference on automotive user interfaces and interactive vehicular applications, pp 1–13
Wu X, Lai IKW (2021) The acceptance of augmented reality tour app for promoting film-induced tourism: the effect of celebrity involvement and personal innovativeness. J Hosp Tour Technol 12(3):454–470. https://doi.org/10.1108/JHTT-03-2020-0054
Xiong J, Hsiang EL, He Z, Zhan T, Wu ST (2021) Augmented reality and virtual reality displays: emerging technologies and future perspectives. Light: Sci Appl 10(1):1–30
Xue H, Sharma P, Wild F (2019) User satisfaction in augmented reality-based training using Microsoft HoloLens. Computers 8(1):9. https://doi.org/10.3390/computers8010009

Yeung AWK, Tosevska A, Klager E, Eibensteiner F, Laxar D, Stoyanov J, Glisic M, Zeiner S, Kulnik ST, Crutzen R, Kimberger O, Kletecka-Pulker M, Atanasov AG, Willschke H (2021) Virtual and augmented reality applications in medicine: analysis of the scientific literature. J Med Internet Res 23(2):e25499. https://doi.org/10.2196/25499

Zhang C, Lu Y (2021) Study on artificial intelligence: The state of the art and future prospects. J Ind Inf Integr 23:100224

Zierau N, Flock K, Janson A, Söllner M, Leimeister J (2021) The influence of AI-based chatbots and their design on users' trust and information sharing in online loan applications. In: Hawaii international conference on system sciences (HICSS), Koloa (Hawaii), USA

Chapter 7
Artificial Intelligence, Machine Learning and Extended Reality: Potential Problem Solvers for Higher Education Issues

Valentin Kuleto, Larisa Mihoreanu, Daniel G. Dinu, Milena P. Ilić, and Dan Păun

Abstract This chapter investigates the potential of new digital technologies and innovative tools in solving Higher Education issues.Higher Education establishments in both Serbia and Romania have experienced dramatic challenges under the process of building up a resilient educational system defined though sound cross-cutting competences according to market demands. Some changes resulted from the EU's transformational reforms approved, the new technologies applied, and teaching and learning methodologies implemented. The primary examples of technology are Massive Open Online Courses (MOOCs) and e-learning platforms, Artificial Intelligence, Machine Learning, Extended Reality, and Blockchain. Examples of innovative teaching techniques include problem-solving, games, interactive instruction, the development of transversal competencies, and knowledge transfer. The targeted and well-structured approach of the agile management within Higher Education has helped to incorporate some of the educational market needs, outlining the potential to get rid of internal injustices as well. The chapter focuses on the survey method used to look at the opinions of students from the Information Technology School in

V. Kuleto · M. P. Ilić (✉)
Information Technology School ITS-Belgrade; LINK Group Belgrade; Faculty of Contemporary Arts Belgrade, University Business Academy in Novi Sad, Belgrade, Serbia
e-mail: milena.ilic@its.edu.rs

V. Kuleto
e-mail: valentin.kuleto@its.edu.rs

L. Mihoreanu
Faculty of Administration and Public Management, Bucharest University of Economic Studies, Bucharest, Romania
e-mail: larisa.mihoreanu@amp.ase.ro

D. G. Dinu
Business Administration Doctoral School, Bucharest University of Economic Studies, Bucharest, Romania
e-mail: daniel.dinu90@gmail.com

D. Păun
Faculty of Physical Education, Spiru Haret University, Bucharest, Romania
e-mail: ushefs_paun.dan@spiruharet.ro

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_7

Belgrade (Serbia) and the Spiru Haret University in Bucharest (Romania). Students took part in an exploratory survey to find out more about their knowledge of Extended Reality, Machine Learning, and Artificial Intelligence as well as their perspectives regarding the difficulties and potential uses of these tools.

7.1 Introduction

Knowledge is becoming obsolete at an exponential rate. Moreover, as the demand for skills changes, the idea that what one learns at school can last for a lifetime becomes a dangerous myth. The concept of lifelong learning is not new and has been part of the policy discourse for over half a century, but it became mired in rhetoric with few implications for policy and practice.

Notably, digitisation and Artificial Intelligence (AI) are changing the nature of employment and the workplace. We know that digitalisation enables skills development and incentivises workers to upgrade their skill sets. Experts expect the pandemic to accelerate the reallocation of labour between economic sectors with important consequences for up- and reskilling. Social participation and citizenship requirements will boost the need to learn. Many other factors, ranging from tourism to civic inclusion, drive the increased demand for language study and digital skills. From a lifelong learning perspective, the distinction between job-related learning and people's motivation to learn because of social, cultural, or personal considerations is blurring. It will require systems to rethink the timing and sequencing of education and skill development across the lifespan, not just boosting adult learning options. This includes the earliest stages of learning that frame the learner's engagement in subsequent steps. All ages will have more complex learning trajectories, necessitating more advanced support systems. Policies must determine what people learn best at what ages and how to allocate resources accordingly.

The concept of lifelong learning urges the rethinking of learning curricula enclosed into a more appropriate architecture of the educational systems. However, some critical learning happens best early in the life course.

As neuroscience and early childhood education research have established, lessening learning opportunities can create irreversible damage.

Long-duration teaching practices are losing value in favour of automated systems. AI and expert systems are used to analyse real-world scenarios for education to get developed. According to Beetham and Sharpe (2013), a sustainable e-learning environment encompasses elements such as e-learning technology and applications, e-teaching concepts, and sustainable development that can be incorporated into the programs, curricula, and pedagogy framework, models, beyond the general theories of intelligent learning or e-learning. The process, lead under the support of AI and expert systems can help choose and suggest a tutoring approach or methodology notably matched to each student's profile (Samah et al. 2011).

Personalisation emphasises academics and self-development. Online courses and online communities help learners develop skills and knowledge. A setting for affordable tutoring and a personalised learner profile (Samah et al. 2011) are examples of a tailored or personalised learning environment.

Teaching by using AI-powered tools in the classroom helps students understand AI technologies, value all their aspects and embrace all AI aspects and connections with education. Additionally, it helps citizens comprehend the enormous potential influence that AI could have on human life (UNESCO 2021).

Artificial intelligence studies include Natural Language Processing (NLP), Machine Learning (ML), and algorithm creation. Applications of AI are transforming educational tools. A few educational uses of AI are personalised learning platforms, automated assessment systems, and facial recognition systems. However, despite the potential advantages of these technologies, the ethical and societal drawbacks of AI systems are rarely effectively discussed in educational settings (Krueger 2017).

The development of artificial intelligence has sped up efforts to build intelligent machines that emulate human behaviour. There have been a huge number of definitions and explanations of what constitutes an AI system as a result of recent advancements in computer science. The use of artificial intelligence can be generally described as "the ability of a digital computer (…) to perform tasks commonly associated with intelligent beings" (Copeland 2020). Using this definition, emphasises the imitation of human behaviour and consciousness, artificial intelligence is "the combination of cognitive automation, machine learning and reasoning". To these, the hypothesis generation and analysis is added together with "the intentional algorithm mutation" (Dieck tom and Jung 2019).

Artificial intelligence relies heavily on algorithms. Therefore, creating sophisticated and evolutionary algorithms is intimately related to the evolution of AI. Algorithms are guidelines or instructions that computers employ to solve issues. Algorithms, as set of rules, are the fundamental building blocks to be followed in different calculations, problem-solving operations, or different computer programs. The computer must adhere to hundreds of lines of code representing mathematical instructions to solve the desired problems (as in computing numerical calculations or processing images and grammar-checking in an essay). For instance, AI algorithms can improve visual perception, learning, speech and face recognition, and other functions. Algorithms can therefore be used to direct the creation of nearly any AI system or application (Borenstein and Howard 2020).

Data and algorithms are used to perform tasks that would usually be done by humans using machine learning. A subfield of computer science called "machine learning" focuses on getting machines to act without being explicitly told to. Machine learning is "a subset of artificial intelligence that studies how computers generate abstract models from data" (Cheney-Lippold 2017, p. 46). The exposure of a learning model to any high-quality amount of data represents the primary machine learning process (Akgun and Greenhow 2021). Machine-learning algorithms first analyse the data to develop a model and forecast future values. There are three steps to machine learning, in other words. There is a first step in which it analyses and gathers the data; It then develops a model to excel at various activities before acting and achieving the intended outcomes without human intervention. A wide variety of AI applications, such as recommender systems and facial recognition systems, are made possible by machine learning (Cath et al. 2018).

Indeed, AI is entirely in line with Computational Thinking (CT) concepts, techniques, and viewpoints since it strives to create robots that can solve issues that, if handled by a human. Designing and executing curricular components should be guided by the goal of demystifying AI and creating informed citizens who can deal with its foundations and effects. Even while programming must not be associated with CT, data suggests that coding is a potent tool for enhancing CT abilities (Lye and Koh 2014; Moreno et al. 2016), especially when employing “low floor and high ceiling” (Resnick et al. 2009) programming environments like Scratch or MIT App Inventor. Tools that attempt to introduce content relevant to AI are required and should adhere to the same design principle: to allow even very young children to participate with little prior knowledge. In this aspect, ML4K (Machine Learning for Kids) is a trailblazer. Four core ML concepts—model, training dataset, learning, and confidence—have been revealed through a practical experience designed to inspire teachers.

Universities may aim to develop CT as a transversal skill for current students and future specialists. Computational Thinking, a technique for representing and solving problems with the use of computers, may be developed using AI based on ML and Extended Reality (XR) applications and games. AI may be applicable and exciting for enhancing CT skills (Garcia et al. 2019). AI is constantly developing, taking over the domain of computer science, and becoming an exciting subject to motivate students’ work. Also, AI issue-solving requires applying several elements, including logical reasoning, critical and analytical thinking, decision-making, and creativity, which align with CT abilities.

ML, which includes a family of algorithms and approaches able to create categorising and predicting models from existing data, has largely been responsible for the rise in popularity and power in recent years. The metaphorical use of the word “learning” discusses the sighting of rules through data analysis. These algorithms fall under one of the categories of reinforcement learning, unsupervised learning, or supervised learning. Supervised learning is used to tackle the majority of the problems that modern AI is quite good at, including, but not limited to, speech recognition, text categorisation, image recognition, and language translation. Data that has been accurately categorised or whose class or expected output is known in advance must be fed to algorithms of this type, called the training set (Garcia et al. 2019).

Innovative teaching resources utilising Extended Reality (XR) have been created. For instance, a half-torso manikin in a virtual reality environment and Basic Life Support (BLS) education data that comply with the 2020 American Heart Association guidelines were used to construct an XR essential life support (XR-BLS) simulator. Chest compressions, ventilation, and an automatic external defibrillator were performed using a head-mounted display (HMD) and hand-tracking equipment in a virtual environment. The XR simulator was simple to use, and the dialogue with the AI instructor was transparent and easy to follow. The XR-BLS simulator is helpful because it can deliver BLS training without calling for teachers and students to congregate (Lee et al. 2022).

The current analysis is based on desk research undertaken by using available academic units of knowledge and experience. It is an exploratory study examining both knowledge and attitudes of 100 Serbian students from the Information Technology School of Belgrade and the Spiru Haret University of Bucharest about AI, ML, and XR in Higher Education Institutions (HEI). The research focus targets the advantages of and links between AI, ML and XR in HEIs, and the obstacles to their common implementation.

7.2 Methodology

7.2.1 Research Aims and Survey

This study directs the identification of the most significant benefits and drawbacks of current technology (XR, AI, ML) in the context of the Romanian and Serbian education sector. For this, an open-ended Likert-scaled survey was created, and the results underwent qualitative analysis. Both Romanian and Serbian students considered completed surveys based on their prior college experiences (Bozkurt et al. 2021; Wang et al. 2019). Unfortunately, only 100 of the replies were confirmed. This research is, therefore, exploratory. Therefore, secondary research necessitates a more thorough analysis of a representative sample. It is regrettably that only 100 of the responses were accurate. Consequently, the research is exploratory. As a result, secondary research demands a more in-depth examination of a representative sample. Using SmartPLS Software 3.0, inferential and variable association analyses were carried out on the collected data.

7.2.2 Variables and Hypothesis

Each of the three categories—AXR (appliances offered by XR in HEI), DXR (difficulties offered by XR in HEI), and AMLAI (appliances offered by AI and ML in HEI)—has nine elements, as presented in Table 7.1 and Fig. 7.1.

The research hypotheses are as follows:

- H1: The innovative tools generate by the implementation of AI and ML in HEIs have a substantial impact on the AI and ML appliances.
- H2: There is a significant correlation between XR performance issues in HEIs and implementation issues with AI and ML in HEIs.
- H3: The devices correlated to the XR performances in HEIs are not significantly impacted by the challenges of XR implementation in HEIs.

Table 7.1 Variable analysed

Code	Variable name
AMLAI1	Personalised learning benefits from AI and machine learning
AMLAI2	The necessity for cross-disciplinary abilities among students is made more accessible by AI and ML
AMLAI3	In the HEI, AI and ML help foster a collaborative learning environment
AMLAI4	Lifelong interaction with alums is encouraged by AI and ML
AMLAI5	AI and ML can be utilised to address security challenges in institutions
AMLAI6	The institution can apply AI and ML to increase its efficiency
AMLAI7	AI and ML allow for the sharing and storing a large amount of data and provide a suitable computing environment
AMLAI8	Researchers are given a good research environment thanks to AI and ML
AMLAI9	AI and ML implementation in HEIs
DAIML1	Cost of AI and ML and Technology Implementation Challenge
DAIML2	Digital competence
DAIML3	A lack of subject-matter experts
DAIML4	AI and machine learning are advantageous for individualised instruction
DAIML5	AI and ML facilitate the need for students to have cross-disciplinary skills
DAIML6	Improper application strategy
AXR1	XR is a valuable tool for aiding education. XR is a valuable tool for streamlining educational procedures
AXR2	Students can actively manage their learning strategies with XR
AXR3	XR inspires and involves pupils
AXR4	XR enables students to learn complex material while having fun simultaneously
AXR5	Through a realistic virtual environment, XR enables students to learn by doing
AXR6	XR helps professors and students interact with one another
AXR7	XR has a beneficial effect on encouraging kids' creative thinking
AXR8	XR encourages pupils' long-term learning and problem-solving efforts
AXR9	The expense of creating course materials is rather significant
DXR1	Technology and equipment are constantly evolving, which compels HEIs to make new expenditures
DXR2	Data in the XR environment is not protected
DXR3	Revisions to the framework's installation requirements for protocols
DXR4	Requirements of protocols installation within the framework, revised
DXR5	Software and hardware for XR were explicitly created for commercial and industrial use, not for educational purposes
DXR6	As a result of technical, economic, and pedagogical limitations, XR still has limited scalability and sustainability

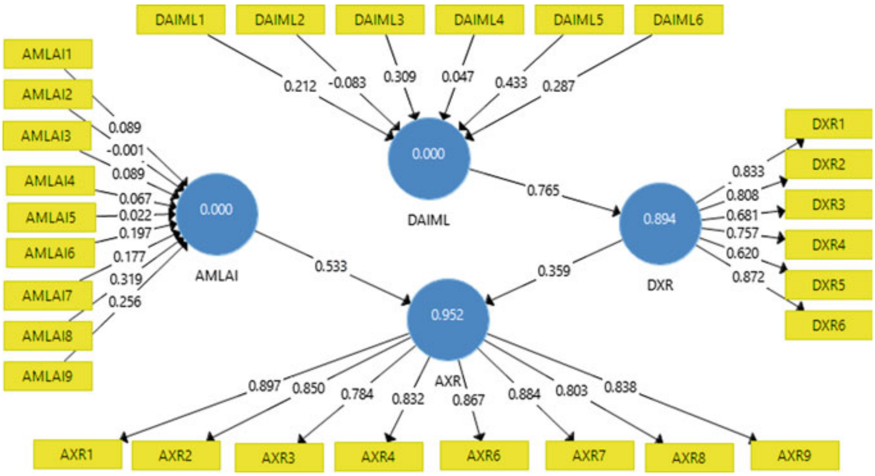


Fig. 7.1 Cronbach Alpha coefficients and path analysis

7.3 Research Results

7.3.1 Construct Reliability and Validity

The SmartPLS software provided us with a wide range of tests for interpreting and assuming the results of our research. As a result, we performed the validation procedures in Table 7.2 to determine the model's consistency (Ray and Saeed 2018; Ilić et al. 2021). The variables in question show unusually high values for composite reliability, Cronbach Alpha and rho A (more than 0.5, the minimum acceptable threshold).

Additionally, it is verified that the latent reflexive variables' AVE (mean extracted Variance) for DXR is 0.58 and 0.71 for AXR. As a result, all our hypotheses have been accepted to varying degrees, as presented by Table 7.2 and Fig. 7.1.

The choice of survey questions is confirmed by the fact that the components (items) for DXR and AXR have very high values (higher than 0.7), as shown in Table 7.2.

Table 7.2 Validation steps/tests

Variable	Cronbach's Alpha	Rho_A	Composite reliability	Average variance extracted
DAIML		1		
DXR	0.894	0.902	0.894	0.588
AMLAI		1		
AXR	0.952	0.953	0.952	0.714

7.3.2 Discriminant Validity

The Heterotrait-Monotrait and the Fornell-Larcker criteria being satisfied, they make the model statistically significant (Fig. 7.2): AMLAI → AXR (0.697), DAIML → DXR (0.765), and DXR → AXR (0.602). Between DAIML and AXR it has been noticed another indirect effect; its influence is quite small (0.275).

The indirect impacts of the model's DAIML → XR → AXR have a low value of 0.275. According to Tables 7.2, 7.3, and 7.4, data proves a positive correlation between the construct indicators AMLAI, DAIML, AXR, and DXR. Strongly correlated with each other are DAIML and DXR (0.765), AMLAI and AXR (0.602), while only a mild positive connection exists between DAIML and AXR (0.697) and (0.548). Compared to the saturated model, our model's estimated Chi-Square is higher (748.466) and (747.629). This proves that the present model is correct, and the H1–H3 hypotheses are also correct (Table 7.4).

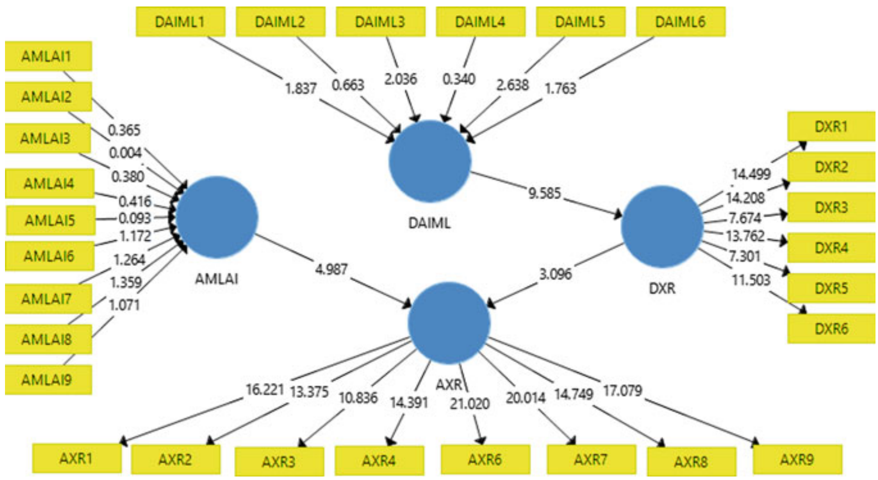


Fig. 7.2 Bootstrapping significance and path coefficients

Table 7.3 Discriminant validity

	Fornell-Larcker criterion				Heterotrait-Monotrait
Variable	DAIML	DXR	AMLAI	AXR	DXR
DAIML					
DXR	0.765	0.767			
AMLAI	0.481	0.455			
AXR	0.548	0.602	0.697	0.845	0.598

Table 7.4 Variable correlation and model fit

	Latent variable correlation				R square	R square adjusted
Variable	DAIML	DXR	AMLAI	AXR		
DAIML	1					
DXR	0.765	1			0.585	0.581
AMLAI	0.481	0.455	1			
AXR	0.548	0.602	0.697	1	0.588	0.579
				Fit summary		
				Saturated model		Estimated model
				747.629		748.466

7.3.3 *Collinearity Statistics VIF*

To assess the usefulness of the variables, the Variance Inflation Factor (VIF) of each construct was calculated using 5000 samples and a 95% bootstrapping method using SmartPLS software. Figure 7.2 displays an overview of the research findings. The results of the bootstrapping test show that multi-collinearity between variables is not present in the overall VIF because the *P*-Values are less than 0.01. In order to get these conclusions, AXR5 was taken out of the model. As a result, it has a VIF more prominent than 5, which has an adverse effect on the model. Therefore, all three hypotheses—H1, H2, and H3—are accepted, with H3 receiving special attention because it states that the route impact has a minor mitigating effect on AXR (2.918).

7.4 Discussion

This study examines the effects of XR, AI, and ML on the Serbian and Romanian educational systems. The ultimate result was a survey with multiple-choice and open-ended questions that could be thoroughly and qualitatively examined. The survey, given to existing and prospective students, was based on earlier research (Götz et al. 2010; CELORR 2019; Ganz and Lecon 2021; Kuleto et al. 2021; Ilić et al. 2021) developed at Spiru Haret University (Bucharest) and ITS (Belgrade). The final results confirm 100 responses as validated. Alternatively, this is a probe, as the name suggests. It is, therefore, impossible to use this study's results as a basis for a comprehensive analysis of the entire statistical population (all Serbian or Romanian students).

SmartPLS Software 3.0 was used to analyse the data for variable association and inferential purposes. The four factors under consideration were AMLAI (appliances given by AI and machine learning in higher education institutions) (and appliances provided by AI and machine learning in higher education institutions). Nine items comprise the formative variable DAIML (difficulties associated with AI and machine learning in higher education institutions) AXR and a six-item constructive variable (appliances provided by XR in higher education institutions). The components of this reflexive variable are eight and DXR (difficulties associated with XR).

We could examine and extrapolate the results from our research with the aid of various tests that SmartPLS offered. AMLAI (appliances supplied by AI and machine learning in higher education institutions) was one of the four variables considered (and appliances provided by AI and machine learning in higher education institutions). DAIML (difficulties associated with AI and machine learning in higher education institutions) AXR has a nine-item formative variable and a six-item constructive variable (appliances provided by XR in higher education institutions). This reflexive variable has eight and DXR as its components (difficulties associated with XR).

With the help of the different tests that SmartPLS provided, we were able to analyze and extrapolate the findings from our research. Table 7.2 contains validation processes that prove the model is reliable and consistent (Hair et al. 2019). Average Variance Extracted, Cronbach Alpha, rho A, and compound reliability are all over 0.7. Due to the convergent validity between the latent reflexive variables (DXR and AXR), the extracted variances (AVE) of 0.58 for DXR and 0.71 for AXR, respectively, also prove convergent validity. All of our hypotheses have likely been partially or fully confirmed.

The results of the Cronbach's Alpha analysis demonstrate the proper design of the DXR and AXR factors or items (higher than 0.7). Because it met the Fornell-Larcker and Heterotrait-Monotrait criteria, the model is statistically significant (Fig. 7.2). DAIML to DXR conversion of 0.765; AMLAI to DXR conversion of 0.797 (0.602). A correlation between DAIML and AXR is also present (0.275).

The results' verification and the validation procedure are aided by the positive connection between AMLAI, DAIML, AXR, and DXR. Strong correlations exist between DXR and AMLAI (0.797), DAIML and AXR (0.765), and DXR and DAIML and AXR (0.797), while only modest correlations exist between DAIML and AXR (0.797), (0.602), (0.548). Based on our model's estimated Chi-Square (748.466) and the saturated model's Chi-Square (747.629), hypotheses H1, H2, and H3 can be accepted (Table 7.4).

The Variance Inflation Factor (VIF) for each construct was estimated using SmartPLS software utilizing 5000 samples and a 95 percent bootstrap method. Figures 7.1 and 7.2 provide a summary of the study's findings. VIF excludes multicollinearity as a result of bootstrapping. AXR5 was not incorporated into the model in order to get these outcomes. Therefore, the model's VIF above five will likely have an undesirable consequence. However, all of these factors indicate that all hypothesis (H1, H2, H3) are true, therefore the moderating effect of DXR on AXR is not particularly substantial—2.918. In HEIs, the AI and ML have some benefits, including giving students a collaborative learning environment, increasing the institution's efficiency, making it easier to share and deposit a large amount of data, and giving researchers a suitable computing environment to conduct their research. Many universities have already begun using this cutting-edge teaching tool, which has been dubbed XR, to help students learn more deeply, become more engaged in the material they are being taught, and have a more enjoyable time doing it (Bucea-Manea-Tonis et al. 2020; Brudermann et al. 2019; Cabero-Almenara et al. 2019; CELORR 2019; Ganz and Lecon 2021; Götz et al. 2010; Hair et al. 2019; Kantsedal et al. 2019; Kuleto et al. 2021; Marujo and Casais 2021). As a result of the high level of interactivity with technology, students can actively manage their learning strategies, allowing them to become more engaged and motivated in their education. Thanks to XR, students can study more thoroughly and sustainably in a virtual world that seems realistic and, as a result, their creative thinking is sustained and encouraged. Additionally, when teaching a new subject, XR emboldens students to interact and connect with their teachers.

The integration and expensive cost of AIML technology are currently impeding its acceptance in higher education institutions, claims H2 (DAIML → DXR). As a result, a digitally proficient expert and an application strategy are required for the implementation of AI and ML systems in higher education institutions. As a result, the adoption of XR in HEIs is positively impacted by AI and ML. However, due to technical, financial, and educational difficulties, poor installation techniques, and a constant need for new equipment, scalability and sustainability are additional issues.

The H3 (DXR → AXR) emphasizes the challenges in deploying XR in HEIs, such as the continuing software and hardware modification requirements and limited sustainability and scalability (due to a lack of prior experience in the field). However, it has little impact on the tools that HEIs can offer, such as the growth and transfer of cross-disciplinary skills suitable for recently graduated students. Other tools make it simpler for students and teachers to work on more varied research projects using collaborative, interactive, and creative learning methods. (Mahmud et al. 2020; Kuleto et al. 2021; Ilić et al. 2021; OECD 2021).

7.5 Conclusion

Information technology is part of our everyday lives, and its uses vary from generation to generation and person to individual. The way we study, live, organise, and work has changed due to new technologies. AI will significantly accelerate this tendency. As the examined works suggest, more efforts should be made to integrate AI-related content into education. To do this, it makes sense to think of AI as a further resource for improving CT skills. “Digital natives”—students born and raised in a world where digital technology permeates every aspect of their lives—make up the majority of today’s student body. Therefore, a university culture shift is necessary to utilise information technology fully.

Digital technologies provide various opportunities and ways to help the transformation of education, which calls for a comprehensive vision of technology-based education, the adjustments needed, and an awareness of its implications. The most pertinent factors that have been put out in the literature to determine how IT affects student performance were examined in this study. Using our data, we were able to investigate a variety of IT-related consequences on student performance.

This study is ground-breaking in that it shows that IT-supported activities, such as student collaboration and interactive learning, positively impact student success. In addition, we found that students with a high level of digital skills performed better and had a higher chance of getting good grades. Teachers' ability to provide timely assistance is a significant factor in this effect.

Both the Serbian and Romanian educational systems are currently going through considerable transformations. Serbia has declared and completed some reforms as part of its continuing EU accession ambitions. The learning results of Serbian students have, however, remained stable over the past few years, with only a few obvious improvements among the best pupils in the country. Many children are unable to demonstrate the core abilities they will need to succeed in school and later on life as a result of this widening educational gap. Additionally, many individuals lack access to socioeconomic groups and areas that hinder the country's development. An overview of the educational systems in Serbia and Romania is developed in this research paper. Additionally, details are supplied on how the assessment and evaluation practices of these countries might be used to improve learning outcomes for all children.

Every research has certain limitations, including this one. Social desirability, harmonisation need, and murky or unreliable metrics are some of the current research's distortions and limits. Therefore, the investigation also emphasizes the need for further research meant to focus on other hidden aspects of the educational systems considered.

References

Akgun S, Greenhow C (2021) Artificial intelligence in education: addressing ethical challenges in K-12 settings. AI Ethics 1–10. Advanced online publication. https://doi.org/10.1007/s43681-021-00096-7

Beetham H, Sharpe R (eds) (2013) Rethinking pedagogy for a digital age: designing for 21st century learning. Rutledge, New York

Borenstein J, Howard A (2020) Emerging challenges in AI and the need for AI ethics education. AI Ethics 1:61–65. https://doi.org/10.1007/s43681-020-00002-7

Bozkurt A, Karadeniz A, Baneres D, Guerrero-Roldán AE, Rodriguez ME (2021) Artificial intelligence and reflections from educational landscape: a review of AI studies in half a century. Sustainability 13(2):800. https://doi.org/10.3390/su13020800/

Brudermann T, Aschemann R, Füllsac M, Posch A (2019) Education for sustainable development 4.0: lessons learned from the University of Graz, Austria. Sustainability 11(8):2347. https://doi.org/10.3390/su11082347

Bucea-Manea-Tonis R, Bucea-Manea-Tonis R, Simion VE, Braicu C, Ilic D, Manea N (2020) Sustainability in higher education: the relationship between work-life balance and XR E-learning facilities. Sustainability 14(12). https://doi.org/10.3390/SU12145872

Cabero-Almenara J, Osuna JB, Llorente-Cejudo C, Martínez MDMF (2019) Educational uses of augmented reality (AR): experiences in educational science. Sustainability 11(18):4990. https://doi.org/10.3390/su11184990

Cath C, Wachter S, Mittelstadt BD, Taddeo M, Floridi L (2018) Artificial intelligence and the 'good society': the US, EU, and UK approach. Sci Eng Ethics 24:505–528. https://www.semanticscholar.org/paper/Artificial-Intelligence-and-the-%E2%80%98Good-Society%E2%80%99%3A-the-Cath-Wachter/cd95a81b10ebfc7d2f4dbc6751c3de05247a13ed

Cheney-Lippold J (2017) We are data: algorithms and the making of our digital selves. NYU Press. https://doi.org/10.2307/j.ctt1gk0941

Copeland B (2020) Artificial intelligence. Encyclopedia Britannica. https://www.britannica.com/technology/artificial-intelligence

Dieck tom MC, Jung T (2019) Augmented reality and virtual reality: the power of AR and VR for business. Springer Nature Switzerland AG. E-book https://link.springer.com/book/https://doi.org/10.1007/978-3-030-06246-0?noAccess=true. ISBN: 978-3-030-06246-0

Ganz N, Lecon C (2021) Opportunities and challenges of XR techniques in HE. In Gómez Chova L, López Martínez A, Candel Torres I, IATED Academy (eds) EDULEARN20 Proceedings of 12th international conference on education and new learning technologies. Online Conference, 6–7 July. IATED, pp 481–490, ISBN: 978-84-09-17979-4, ISSN: 2340-1117. https://doi.org/10.21125/edulearn.2020.0207. https://library.iated.org/view/GANZ2020OPP

Garcia JDR, Leon JM, Gonzalez MR, Robles G (2019) Developing computational thinking at school with machine learning: an exploration. In: Book Group Author: IEEE 2019 International symposium on computers in education (SIIE), Book series international symposium on computers in education

Götz O, Liehr-Gobbers K, Krafft M (2010) Evaluation of structural equation models using the partial least squares (PLS) approach. In: Esposito Vinzi V, Chin W, Henseler J Wang (eds) Handbook of partial least squares. Springer handbooks of computational statistics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-32827-8_30

Hair JF, Risher JJ, Sarstedt M, Ringle CM (2019) When to use and how to report the results of PLS-SEM. Eur Bus Rev 31(1):2–24. https://doi.org/10.1108/EBR-11-2018-0203

Ilić MP, Păun D, Popović Šević N, Hadžić A, Jianu A (2021) Needs and performance analysis for changes in higher education and implementation of artificial intelligence, machine learning, and extended reality. Educ Sci 11(10):568. https://doi.org/10.3390/educsci11100568

Kantsedal N, Ponomarenko O, Dorohan'-Pisarenko L, Liaska O (2019) The methods of using interactive technologies during teaching foundations of scientific research at higher educational establishments. Indep J Manag Prod 10:778–797

Krueger N (2017) Artificial intelligence has infiltrated our lives. Can it improve learning? Int Soc Technol Educ (ISTE) 17 June. https://www.iste.org/explore/Empowered-Learner/Artificial-intelligence-has-infiltrated-our-lives.-Can-it-improve-learning%3F

Kuleto V, Ilić MP, Stanescu M, Ranković M, Šević NP, Păun D, Teodorescu S (2021) Extended reality in higher education, a responsible innovation approach for Generation Y and Generation Z. Sustainability 13(21):11814. https://doi.org/10.3390/su132111814

Lee DK, Choi H, Jheon S, Jo YH, Im CW Il SY (2022) Development of an extended reality simulator for basic life support training, 10. https://doi.org/10.1109/JTEHM.2022.3152365

Lye SY, Koh JHL (2014) Review on teaching and learning of computational thinking through programming: what is next for K12? Comput Hum Beh 41:51–61

Mahmud SND, Husnin H, Tuan Soh TM (2020) Teaching presence in online gamified education for sustainability learning. Sustainability 12(9):3801. https://doi.org/10.3390/su12093801

Marujo HA, Casais M (2021) Educating for public happiness and global peace: contributions from a Portuguese UNESCO Chair towards the sustainable development goals. Sustainability 13(16):9418. https://doi.org/10.3390/su13169418

Moreno JL, Robles G, Román-González M (2016) Code to learn: where does it belong in the K-12 curriculum? JITE: Res 15:283–303

OECD (2021) Building the future of education. https://www.oecd.org/education/future-of-education-brochure.pdf. Accessed 04.06.2022

Ray S, Saeed M (2018) Applications of educational data mining and learning analytics tools in handling Big Data in higher education. In: Alani M, Tawfik H, Saeed M, Anya O (eds)

Applications of Big Data analytics. Springer, Cham. https://doi.org/10.1007/978-3-319-76472-6_7

Resnick JM, Monroy-Hernández A, Rusk N, Eastmond E, Brennan K, Millner A, Rosenbaum E, Silver J, Silverman B, Kafai Y (2009) Scratch: programming for all. Commun ACM 52(11):60–67. https://doi.org/10.1145/1592761.1592779

Samah NA, Yahaya N, Ali MB (2011) Individual differences in the online personalised learning environment. Educ Res Rev 6(7):516–521. https://academicjournals.org/article/article1379765314_Samah%20et%20al.pdf

Smith H (2020) Algorithmic bias: should students pay the price? AI Soc 35:1077–1078. https://doi.org/10.1007/s00146-020-01054-3

UNESCO Fengchun M, Holmes W, Ronghuai H, Hui Z (2021) AI and education: guidance for policymakers. ISBN: 978-92-3-100447-6. CC BY-SA 3.0 IGO [10451]. https://en.unesco.org/artificial-intelligence/education

University of Waterloo (2019) Centre for extended learning open resource repository—CELORR. Extended reality. https://contensis.uwaterloo.ca/sites/open/resources/CEL-ORR/toc/modules/extended-reality.aspx

Wang Y, Hong S, Tai C (2019) China's efforts to lead the way in AI start in its classrooms. Wall Street J. 24.10.2019. https://www.wsj.com/articles/chinas-efforts-to-lead-the-way-in-ai-start-in-its-classrooms-11571958181

Chapter 8
Augmented Reality and Artificial Intelligence in Education: Toward Immersive Intelligent Tutoring Systems

Georgios Lampropoulos

Abstract The integration of Artificial Intelligence (AI) and Augmented Reality (AR) in education as a means to address the new educational needs and to provide high quality education is gaining ground. AR provides engaging, interactive and motivating mixed reality environments while AI offers personalized experiences. The aim of this chapter is to explore how the convergence of AI and AR can impact the educational domain and lead to the development of more interactive and immersive Intelligent Tutoring Systems (ITS). This chapter presents the technologies of AI, AR and ITS and examines their use in educational settings. Then, it showcases the findings of recent related studies, discusses the potentials and merits that their combination engenders in educational contexts and provides future research directions. Based on the findings, it can be said that the convergence of AI and AR can result in the development of interactive, immersive, engaging and personalized learning experiences. These experiences can be applied in all educational levels and have the potential to take into account learners' psychological, cognitive and motivational states as well as their unique characteristics, knowledge, preferences, interests and performance. These experiences can help to meet the educational demand, positively impact teaching and learning activities and increase learning outcomes and motivation.

8.1 Introduction

The COVID-19 pandemic has highlighted the drastic need for ubiquitous learning (Daniel 2020). Modern students are accustomed to handling digital devices and media on a daily basis (Henderson et al. 2015). This has drastically affected the way of their learning and getting informed (Lampropoulos et al. 2021b; Prensky

G. Lampropoulos (✉)
Department of Information and Electronic Engineering, International Hellenic University, Thessaloniki, Greece
e-mail: lamprop.geo@gmail.com

School of Humanities, Hellenic Open University, Patras, Greece

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_8

2001). Thus, students are seeking for more engaging learning experiences which take into account their needs, preferences, existing knowledge, performance and unique traits (Anastasiadis et al. 2018; Robinson et al. 2013; Zeidler and Nichols 2009). In both face-to-face and distance learning, students require more personalized learning which, in most cases, is not possible to provide due to the number of students in traditional learning and due to distance and time differences in online learning (Xie et al. 2019; Zhang et al. 2020). As a result, the need for the transformation of education and the appropriate management to ensure its success is evident (Abad-Segura et al. 2020). Following a student-centered approach, the integration of Information and Communications Technologies (ICT) and novel technologies in education can correspond to the new educational requirements, facilitate teachers, provide education of high quality, offer inclusive learning opportunities and assist the transition to technology-enhanced learning (López-Belmonte et al. 2022).

Furthermore, adaptive, immersive and personalized learning environments and experiences can increase students' learning achievements, motivation, engagement and knowledge acquisition (Lampropoulos et al. 2022). Adaptive learning provides instantaneous feedback, offers interactive experiences and creates dynamic and interactive content in an automated manner which can be modified based on learners' comprehension, performance and individual differences (Hattie 2008; Kerr 2016; Martin et al. 2020). Immersive learning utilizes Extended Reality (XR) technologies to engulf learners in engaging and interactive environments while providing them with a sense of presence (De Freitas et al. 2010; Herrington et al. 2007). Personalized learning is an approach which focuses on providing learning experiences that are tailored and customized to the needs of each learner and assists them in achieving their goals and enriching their knowledge and skills (Bernacki et al. 2021; Shemshack and Spector 2020). The combination of Augmented Reality (AR) with Artificial Intelligence (AI) can lead to the development of advanced Intelligent Tutoring Systems (ITS) which generate the conditions required for such environments to be created. Additionally, these ITS will be characterized by both interactivity and immersion due to the interactive AR objects and by personalization and ubiquity owing to their ability to provide one-to-one customized learning opportunities at any given time and place while taking learners' unique traits and abilities into account. Consequently, the aim of this study is to explore how the convergence of AR and AI can impact education through the development of immersive and interactive ITS.

8.2 Augmented Reality

Due to its ability to intertwine the virtual with the real environment, AR is closer to the real environment according to the reality-virtuality continuum (Milgram and Kishino 1994) and enables users to interact with virtual objects that co-exist with real-world objects (Zhou et al. 2008). As it allows users to constantly see the real environment, it is differentiated from Virtual Reality (VR) which perceptually surrounds and immerses users in a completely virtual environment (Burdea and Coiffet 2003;

Lampropoulos et al. 2021a). AR aims at enriching users' physical environment with virtual objects (Billinghurst et al. 2015). Particularly, AR involves the use of computer units to create digital objects and information and embed them in the appropriate time and space so that users can perceive them through their senses and interact with them in real time (Azuma 1997; Caudell and Mizell 1992; Lampropoulos et al. 2020).

AR can be effectively used to enrich both learning and teaching activities in all educational levels (López-Belmonte et al. 2019). Due to its immersive and interactive nature, AR can create learning experiences that lead to increased learners' academic achievements, learning motivation, active involvement and knowledge gain (Avila-Garzon et al. 2021; Garzón et al. 2019; Lampropoulos et al. 2022). Additionally, it provides the required conditions in which inclusive education can flourish (Pozo-Sánchez et al. 2022). When combined with artificial intelligence and learning analytics, AR has the potential to provide personalized, ubiquitous and adaptive learning experiences.

8.3 Artificial Intelligence

AI can be used in combination with other state-of-the-art technologies to create autonomous, sophisticated and rational decision-making systems that do not require any human intervention (Duan et al. 2019). AI is associated with the capabilities of a system to interpret and learn from external data sources to achieve particular aims in a flexible and adaptable manner (Haenlein and Kaplan 2019). AI is inspired by the way humans use their nervous system to feel, think and reason and constitutes a scientific field as well as a set of computational technologies (Stone et al. 2016). Particularly, AI refers to a system ability to mimic human actions and simulate human intelligence to gain enhanced perception, learning and reasoning capabilities, autonomy, rationality and adaptability (Brynjolfsson and Mcafee 2017; Lampropoulos 2022a; Li and Du 2017; Stone et al. 2016). It can be categorized into artificial narrow intelligence, artificial general intelligence and artificial superintelligence when taking into account the evolutionary stages of AI and classified into humanized, human-inspired and analytical AI when considering its intelligence types (Russell 2010a, b). AI can yield numerous benefits to several domains (Bughin et al. 2017).

AI has the potential to enrich and transform the educational sector (Holmes et al. 2020), bring about new teaching and learning methods (Pedro et al. 2019) and foster the redesign of cognition, knowledge and culture (Hwang et al. 2020). Recent systematic literature review studies have explored the use and impact of AI in education and presented its potential applications, merits and challenges (Chen et al. 2020a, b; Chen et al. 2022; Tang et al. 2021; Zhai et al. 2021). The integration of AI in education can empower learners' agency, improve personalization, increase sustainability, allow customized feedback and lead to data-driven, student-centered and personalized learning experiences (Chiu and Chai 2020; Lampropoulos 2022b; Ouyang and Jiao 2021). Nonetheless, there still remain several theory and application gaps that need to be explored to effectively adopt and use AI in education (Chen et al. 2020a, b).

8.4 Intelligent Tutoring Systems

The use of AI in education is becoming increasingly popular with emphasis being put on creating ITS that provide adaptive and personalized learning experiences by customizing the instructional strategies and activities according to each learner (Keleş et al. 2009; Mousavinasab et al. 2021). The main aim of ITS is to provide individualized learning that surpass that of a sophisticated human tutor and to adapt to learners at a fine-grained level while taking learners' unique traits, characteristics, special needs, learning pace and performance into consideration (Anderson et al. 1985; Sleeman and Brown 1982). Particularly, ITS refer to computer-assisted learning environments that use computational models developed in other related scientific fields (e.g., learning sciences, cognitive sciences, mathematics, etc.) and can be applied in various subjects and educational levels to assist learners in gaining domain-specific, cognitive and metacognitive knowledge (Ma et al. 2014; Steenbergen-Hu and Cooper 2014). Additionally, ITS can be regarded as interactive, flexible, student-centered and self-paced learning environments (Graesser et al. 2018). They are able to identify and track learners' psychological states and address their needs by adaptively adjusting and responding (Graesser et al. 2012; Shute and Zapata-Rivera 2007).

ITS capitalize on complex computational models which are based on learners' motivational and cognitive states to provide computerized learning environments; hence, they include several components (e.g., modeling, feedback, scaffolding, prediction as well as adaptive lessons and activities), models (e.g., expert, student, domain, pedagogical, user interface) and knowledge types (e.g., knowledge regarding the instructional domain, student and pedagogical strategies) (Conati 2009; Keleş et al. 2009; Xu et al. 2019). Therefore, several factors should be taken into account when designing and developing ITS (Hwang 2003), such as the ability to provide learners with customized and instantaneous feedback (Gross et al. 2015), to manage their affective states (Frasson and Chalfoun 2010a, b) and to opt for the most effective learning style (Crockett et al. 2017).

While focusing on personalized, one-on-one, face-to-face and computer-based tutoring (VanLehn 2006), ITS are almost as efficient as human tutoring systems (VanLehn 2011) and moderately increase student learning (Kulik 1994) in well-defined domains. In ill-defined domains, human tutors still outperform ITS (Gross et al. 2015). The integration of ITS in both K-12 and higher education has proven to be an effective learning method when being used as the main means of instruction delivery or as a supplementary tool (Kulik and Fletcher 2016; Ma et al. 2014;

Steenbergen-Hu and Cooper 2014; Xu et al. 2019). Nonetheless, there is a clear need to further explore mobile-based ITS (Mousavinasab et al. 2021) and address the open challenges of ITS (Conati 2009).

8.5 Discussion

As students grow up with Information and Communication Technologies (ICT) and have constant access to digital devices and new media, the way of acquiring information and knowledge has drastically changed. Particularly, students pursue more meaningful and personalized learning experiences which would satisfy their specific needs. Additionally, students seek for more engaging and motivating learning environments that will allow them to acquire knowledge and hone their skills. To address the existing and upcoming educational demands, educational content is enriched by multimedia elements and technology-enhanced learning is being more widely adopted (Sung et al. 2016). Concurrently, the teaching and learning approaches and methods are constantly evolving (Romero et al. 2015).

AR is an interactive and immersive technology that is being more widely applied in the educational sector. Due to its nature, it creates safe learning environments and experiences that combine the virtual with the real world something that would be impossible otherwise. AR enables students to acquire hands-on experiences, view objects from different perspectives and learn both individually and collaboratively. This novel technology has been proven to be an effective learning tool that not only increases learners' active involvement, communication and learning motivation, but also leads to increased learning outcomes and learners' development (Lampropoulos et al. 2022). Additionally, AR is positively assessed by the public who mostly expresses positive emotions and has a positive attitude toward both its use in general as well as its adoption and integration in education (Lampropoulos et al. 2022c). AR can be used in combination with different technologies and approaches to further enhance its functionality and, in turn, the benefits to be yielded.

AI has been applied in several domains which, in most cases, has led to their transformation and improvement. The educational domain is not an exception. As the amount of data and the computational power increases, so does the effectiveness of AI. ITS use different models and AI methods to provide learners with personalized experiences. Particularly, ITS can identify learners' psychological, cognitive and motivational states and provide individualized learning tasks and feedback based on learners' unique characteristics, existing knowledge, interests, preferences and performance (Sleeman and Brown 1982). Despite their several benefits, ITS are not always presented in a user-friendly and interactive manner.

Through the convergence of AR and AI, the individual drawbacks of AR and ITS can be overcome while simultaneously better outcomes and new learning opportunities arise. By combining these technologies, immersive, interactive and personalized experiences that take place in both digital and physical environments can be created. This combination can lead to engaging and effective learning experiences that are dynamically individualized based on each learner and which offer immediate and targeted instructions and feedback in real time. Systems and applications that fuse AR and AI have the potential to engage students in interactive and sustained reasoning activities based on a deep understanding of their behavior and preferences. Additionally, they can be applied in all levels and contexts in both education and training while providing easy access to high quality education, enabling learning in a meaningful and effective manner and increasing learning achievements.

8.6 Conclusion

The educational domain is undergoing a drastic transformation. Simultaneously, learners are seeking for more personalized and effective learning methods and tools. The adoption and integration of state-of-the-art technologies in the educational sector can bring about several benefits and cope with the challenges that come up. This chapter aimed at examining how the convergence of AR and AI could potentially affect education by providing students with more interactive and immersive ITS. Hence, the chapter went over the related technologies, presented findings of related studies and discussed the potential of their convergence.

When used in education in a student-centered manner, AR can create to more engaging, motivating and enjoyable learning experiences which lead to improved learning outcomes and satisfaction. Through the use of ITS, AI can provide personalized learning experiences by taking learners' characteristics, personality traits, knowledge, performance and preferences into account and address their needs in a flexible and adaptable way. The fusion of AR and AI has the potential to create new learning environments and opportunities that could successfully meet the educational needs, enhance learners' motivation and performance and provide inclusive, individualized and ubiquitous learning in both formal and informal settings and throughout all educational levels. Furthermore, these environments are immersive and interactive in nature and can increase students' engagement and learning outcomes, positively affect teaching and learning activities and meet the existing and upcoming educational needs. Future research will focus on developing such systems and evaluating their effectiveness in cross-cultural and cross-country settings.

Acknowledgements The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the 3rd Call for HFRI PhD Fellowships (Fellowship Number: 6454).

References

Abad-Segura E, González-Zamar MD, Infante-Moro JC, Ruipérez García G (2020) Sustainable management of digital transformation in higher education: global research trends. Sustainability 12(5):2107. https://doi.org/10.3390/su12052107

Anastasiadis T, Lampropoulos G, Siakas K (2018) Digital game-based learning and serious games in education. Int J Adv Sci Res Eng 4(12):139–144. https://doi.org/10.31695/ijasre.2018.33016
Anderson JR, Boyle CF, Reiser BJ (1985) Intelligent tutoring systems. Science 228(4698):456–462. https://doi.org/10.1126/science.228.4698.456
Avila-Garzon C, Bacca-Acosta J, Kinshuk, Duarte J, Betancourt J (2021) Augmented reality in education: an overview of twenty-five years of research. Contemp Educ Technol 13(3):ep302. https://doi.org/10.30935/cedtech/10865
Azuma RT (1997) A survey of augmented reality. Presence: Teleoperators Virtual Environ 6(4):355–385. https://doi.org/10.1162/pres.1997.6.4.355
Bernacki ML, Greene MJ, Lobczowski NG (2021) A systematic review of research on personalized learning: personalized by whom, to what, how and for what purpose(s)? Educ Psychol Rev 33(4):1675–1715. https://doi.org/10.1007/s10648-021-09615-8
Billinghurst M, Clark A, Lee G (2015) A survey of augmented reality. Found Trends Hum-Comput Interact 8(2–3):73–272. https://doi.org/10.1561/1100000049
Brynjolfsson E, Mcafee A (2017) Artificial intelligence, for real. Harv Bus Rev 1:1–31
Bughin J, Hazan E, Ramaswamy S, Chui M, Allas T, Dahlstrom P, Henke N, Trench M (2017) Artificial intelligence: the next digital frontier? McKinsey Global Institute
Burdea GC, Coiffet P (2003) Virtual reality technology. John Wiley and Sons. https://doi.org/10.1162/105474603322955950
Caudell PT, Mizell DW (1992) Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Hawaii international conference on system sciences, 2. https://doi.org/10.1109/HICSS.1992.183317
Chen L, Chen P, Lin Z (2020a) Artificial intelligence in education: a review. IEEE Access 8:75264–75278. https://doi.org/10.1109/access.2020.2988510
Chen X, Xie H, Zou D, Hwang GJ (2020b) Application and theory gaps during the rise of artificial intelligence in education. Comput Educ: Artif Intell 1:100002. https://doi.org/10.1016/j.caeai.2020.100002
Chen X, Zou D, Xie H, Cheng G, Liu C (2022) Two decades of artificial intelligence in education. Educ Technol Soc 25(1):28–47
Chiu TKF, Chai C (2020) Sustainable curriculum planning for artificial intelligence education: a self-determination theory perspective. Sustainability 12(14):5568. https://doi.org/10.3390/su12145568
Conati C (2009) Intelligent tutoring systems: new challenges and directions. In: Twenty-first international joint conference on artificial intelligence
Crockett K, Latham A, Whitton N (2017) On predicting learning styles in conversational intelligent tutoring systems using fuzzy decision trees. Int J Hum Comput Stud 97:98–115. https://doi.org/10.1016/j.ijhcs.2016.08.005
Daniel SJ (2020) Education and the COVID-19 pandemic. Prospects 49(1–2):91–96. https://doi.org/10.1007/s11125-020-09464-3
De Freitas S, Rebolledo-Mendez G, Liarokapis F, Magoulas G, Poulovassilis A (2010) Learning as immersive experiences: using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. Br J Edu Technol 41(1):69–85. https://doi.org/10.1111/j.1467-8535.2009.01024.x
Duan Y, Edwards JS, Dwivedi YK (2019) Artificial intelligence for decision making in the era of big data—evolution, challenges and research agenda. Int J Inf Manage 48:63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021
Frasson C, Chalfoun P (2010) Managing learner's affective states in intelligent tutoring systems. In: Studies in computational intelligence, pp 339–358. https://doi.org/10.1007/978-3-642-14363-2_17
Garzón J, Pavón J, Baldiris S (2019) Systematic review and meta-analysis of augmented reality in educational settings. Virtual Reality 23(4):447–459. https://doi.org/10.1007/s10055-019-00379-9

Graesser AC, Conley MW, Olney A (2012) Intelligent tutoring systems. In: APA educational psychology handbook, vol 3: Application to learning and teaching, pp 451–473. https://doi.org/10.1037/13275-018

Graesser AC, Hu X, Sottilare R (2018) Intelligent tutoring systems. In: International handbook of the learning sciences, pp 246–255. https://doi.org/10.4324/9781315617572-24

Gross S, Mokbel B, Hammer B, Pinkwart N (2015) Learning feedback in intelligent tutoring systems. KI—Künstliche Intelligenz 29(4):413–418. https://doi.org/10.1007/s13218-015-0367-y

Haenlein M, Kaplan A (2019) A brief history of artificial intelligence: on the past, present and future of artificial intelligence. Calif Manage Rev 61(4):5–14. https://doi.org/10.1177/0008125619864925

Hattie J (2008) Visible learning: a synthesis of over 800 meta-analyses relating to achievement. Routledge

Henderson M, Selwyn N, Finger G, Aston R (2015) Students' everyday engagement with digital technology in university: exploring patterns of use and 'usefulness'. J High Educ Policy Manag 37(3):308–319. https://doi.org/10.1080/1360080x.2015.1034424

Herrington J, Reeves TC, Oliver R (2007) Immersive learning technologies: realism and online authentic learning. J Comput High Educ 19(1):80–99. https://doi.org/10.1007/bf03033421

Holmes W, Bialik M, Fadel C (2020) Artificial intelligence in education: promises and implications for teaching and learning. Center for Curriculum Redesign

Hwang GJ (2003) A conceptual map model for developing intelligent tutoring systems. Comput Educ 40(3):217–235. https://doi.org/10.1016/s0360-1315(02)00121-5

Hwang GJ, Xie H, Wah BW, Gašević D (2020) Vision, challenges, roles and research issues of artificial intelligence in education. Comput Educ: Artif Intell 1:100001. https://doi.org/10.1016/j.caeai.2020.100001

Keleş A, Ocak R, Keleş A, Gülcü A (2009) ZOSMAT: web-based intelligent tutoring system for teaching–learning process. Expert Syst Appl 36(2):1229–1239. https://doi.org/10.1016/j.eswa.2007.11.064

Kerr P (2016) Adaptive learning: ELT J 70(1):88–93. https://doi.org/10.1093/elt/ccv055

Kulik JA, Fletcher JD (2016) Effectiveness of intelligent tutoring systems. Rev Educ Res 86(1):42–78. https://doi.org/10.3102/0034654315581420

Kulik JA (1994) Meta-analytic studies of findings on computer-based instruction. In: Technology assessment in education and training. Lawrence Erlbaum Associates, Inc., pp 9–33

Lampropoulos G, Keramopoulos E, Diamantaras K (2020) Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: a review. Vis Inform 4(1):32–42. https://doi.org/10.1016/j.visinf.2020.01.001

Lampropoulos G, Keramopoulos E, Diamantaras K, Evangelidis G (2022) Augmented reality and gamification in education: a systematic literature review of research, applications and empirical studies. Appl Sci 12(13):6809. https://doi.org/10.3390/app12136809

Lampropoulos G, Siakas K, Makkonen P, Siakas E (2021b) A 10-year longitudinal study of social media use in education. Int J Technol Educ, pp 373–398. https://doi.org/10.46328/ijte.123

Lampropoulos G, Barkoukis V, Burden K, Anastasiadis T (2021a) 360-degree video in education: an overview and a comparative social media data analysis of the last decade. Smart Learn Environ 8(1). https://doi.org/10.1186/s40561-021-00165-8

Lampropoulos G (2022a) Artificial intelligence, big data and machine learning in industry 4.0. Encyclopedia of data science and machine learning. IGI Global. https://doi.org/10.4018/978-1-7998-9220-5.ch125

Lampropoulos G (2022b) Educational data mining and learning analytics in the 21st century. Encyclopedia of data science and machine learning. IGI Global. https://doi.org/10.4018/978-1-7998-9220-5.ch098

Lampropoulos G, Keramopoulos E, Diamantaras K, & Evangelidis G (2022c). Augmented reality and virtual reality in education: Public perspectives, sentiments, attitudes, and discourses. Educ Sci, 12(11), 798. https://doi.org/10.3390/educsci12110798

Li D, Du Y (2017) Artificial intelligence with uncertainty. CRC Press. https://doi.org/10.1201/9781315366951

López-Belmonte J, Moreno-Guerrero AJ, López Núñez JA, Pozo Sánchez S (2019) Analysis of the productive, structural and dynamic development of augmented reality in higher education research on the web of science. Appl Sci 9(24):5306. https://doi.org/10.3390/app9245306

López-Belmonte J, Moreno-Guerrero AJ, Marín-Marín JA, Lampropoulos G (2022) The impact of gender on the use of augmented reality and virtual reality in students with ASD. Educ Knowl Soc (EKS) 23. https://doi.org/10.14201/eks.28418

Ma W, Adesope OO, Nesbit JC, Liu Q (2014) Intelligent tutoring systems and learning outcomes: a meta-analysis. J Educ Psychol 106(4):901–918. https://doi.org/10.1037/a0037123

Martin F, Chen Y, Moore RL, Westine CD (2020) Systematic review of adaptive learning research designs, context, strategies and technologies from 2009 to 2018. Educ Tech Res Dev 68(4):1903–1929. https://doi.org/10.1007/s11423-020-09793-2

Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst 77(12):1321–1329

Mousavinasab E, Zarifsanaiey N, Kalhori SRN, Rakhshan M, Keikha L, Ghazi Saeedi M (2021) Intelligent tutoring systems: a systematic review of characteristics, applications and evaluation methods. Interact Learn Environ 29(1):142–163. https://doi.org/10.1080/10494820.2018.1558257

Ouyang F, Jiao P (2021) Artificial intelligence in education: the three paradigms. Comput Educ: Artif Intell 2:100020. https://doi.org/10.1016/j.caeai.2021.100020

Pedro F, Subosa M, Rivas A, Valverde P (2019) Artificial intelligence in education: challenges and opportunities for sustainable development

Pozo-Sánchez S, Lampropoulos G, López-Belmonte J (2022) Comparing gamification models in higher education using face-to-face and virtual escape rooms. J New Approaches Educ Res 11(2):307. https://doi.org/10.7821/naer.2022.7.1025

Prensky M (2001) Digital natives, digital immigrants part 2: do they really think differently? On Horiz 9(6):1–6. https://doi.org/10.1108/10748120110424843

Robinson R, Molenda M, Rezabek L (2013) Facilitating learning. In Educational technology. Routledge, pp 27–60

Romero M, Usart M, Ott M (2015) Can serious games contribute to developing and sustaining 21st century skills? Games Cult 10(2):148–177. https://doi.org/10.1177/1555412014548919

Russell SJ (2010) Artificial intelligence a modern approach. Pearson Education, Inc.

Shemshack A, Spector JM (2020) A systematic literature review of personalized learning terms. Smart Learn Environ 7(1). https://doi.org/10.1186/s40561-020-00140-9

Shute VJ, Zapata-Rivera D (2007) Adaptive technologies. ETS Res Rep Ser, 1–34. https://doi.org/10.1002/j.2333-8504.2007.tb02047.x

Sleeman D, Brown JS (1982) Intelligent tutoring systems. Academic Press, London

Steenbergen-Hu S, Cooper H (2014) A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. J Educ Psychol 106(2):331–347. https://doi.org/10.1037/a0034752

Stone P, Brooks R, Brynjolfsson E, Calo R, Etzioni O, Hager G, Hirschberg J, Kalyanakrishnan S, Kamar E, Kraus S, Leyton-Brown K, Parkes D, Press W, Saxenian A, Shah J, Tambe M, Teller A (2016) Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence

Sung YT, Chang KE, Liu TC (2016) The effects of integrating mobile devices with teaching and learning on students' learning performance: a meta-analysis and research synthesis. Comput Educ 94:252–275. https://doi.org/10.1016/j.compedu.2015.11.008

Tang KY, Chang CY, Hwang GJ (2021) Trends in artificial intelligence-supported e-learning: a systematic review and co-citation network analysis (1998–2019). Interact Learn Environ, 1–19. https://doi.org/10.1080/10494820.2021.1875001

VanLehn K (2006) The behavior of tutoring systems. Int J Artif Intell Educ 16(3):227–265

VanLehn K (2011) The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. Educ Psychol 46(4):197–221. https://doi.org/10.1080/00461520.2011.611369

Xie H, Chu HC, Hwang GJ, Wang CC (2019) Trends and development in technology-enhanced adaptive/personalized learning: a systematic review of journal publications from 2007 to 2017. Comput Educ 140:103599. https://doi.org/10.1016/j.compedu.2019.103599

Xu Z, Wijekumar K, Ramirez G, Hu X, Irey R (2019) The effectiveness of intelligent tutoring systems on k-12 students' reading comprehension: a meta-analysis. Br J Edu Technol 50(6):3119–3137. https://doi.org/10.1111/bjet.12758

Zeidler DL, Nichols BH (2009) Socioscientific issues: theory and practice. J Elem Sci Educ 21(2):49–58. https://doi.org/10.1007/bf03173684

Zhai X, Chu X, Chai CS, Jong MSY, Istenic A, Spector JM, Liu JB, Yuan J, Li Y (2021) A review of artificial intelligence (AI) in education from 2010 to 2020. Complexity 2021:1–18. https://doi.org/10.1155/2021/8812542

Zhang L, Basham JD, Yang S (2020) Understanding the implementation of personalized learning: a research synthesis. Educ Res Rev 31:100339. https://doi.org/10.1016/j.edurev.2020.100339

Zhou F, Duh HBL, Billinghurst M (2008) Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR. In: 2008 7th IEEE/ACM international symposium on mixed and augmented reality. https://doi.org/10.1109/ismar.2008.4637362

Chapter 9
Augmented/Virtual Reality and Artificial Intelligence in Dental Education and Research

Narayan H. Gandedkar, Matthew Wong, Sabarinath Prasad, and M. Ali Darendeliler

Abstract With the advent of 'Information Technology' phenomenon, such as deep learning (DL), machine learning (ML) and artificial neural networks (ANN), dentistry has taken a new turn. Application of information technology in dental education and research, and its interaction with strategic frameworks, professionalism, and patient care is of paramount importance. The role of Virtual Reality (VR), Augmented Reality (AR) and Artificial Intelligence (AI) in improving dental school education and training has been primarily driven to overcome traditional teaching limitations. To overcome these short comings, several Virtual Reality systems have been adopted with an engagement of wide-ranging sensory channels by virtual simulation. This includes Fully immersive VR (e.g., head-mounted displays, HMD's), Semi-immersive VR (e.g., cave automatic virtual environment, CAVE), and non-immersive screen-based VR (Computer-aided design systems, CAD). This chapter provides an insight of current trends and future forecasts of application of VR, AR, and AI in dentistry related education and research.

N. H. Gandedkar (✉) · M. Wong · M. Ali Darendeliler
Department of Orthodontics and Paediatric Dentistry, Faculty of Medicine and Health, School of Dentistry, The University of Sydney, Sydney, Australia
e-mail: narayan.gandedkar@sydney.edu.au

M. Ali Darendeliler
e-mail: ali.darendeliler@sydney.edu.au

S. Prasad
Department of Orthodontics, Hamdan Bin Mohammed College of Dental Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates
e-mail: sabarinath.prasad@mbru.ac.ae

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*, Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_9

9.1 Introduction

Information technology phenomenon, such as deep learning (DL), machine learning (ML) and artificial neural networks (ANN), especially augmented reality (AR) is not only embraced to revolutionise teaching and imparting knowledge in dentistry proceedings. It has also generated profound interest in the education and research area in the form of Artificial Intelligence in Education (AIEd) and Learning Analytics (LA) and Augmented Reality Learning Experience (ARLE) (Radianti et al. 2020; Park and Park 2018; Geroimenko 2020; Drljević et al. 2017). The primary goal of Artificial Intelligence in Education (AIEd) is to facilitate a 'human–AI hybrid' corroborative platform to enhance student learning through applications of AI, human–computer interaction, and the learning sciences (Luckin et al. 2016; Renz et al.), whilst Augmented Reality Learning Experience (ARLE) comprises of overlay of real world with virtual elements to empower the educator to deliver appropriate, persuasive, and immersive experiences (Drljević et al. 2017; Dunleavy et al. 2009).

In the current era of rapid information transformation, a pertinent question arises—does dental education and research require a curriculum overhaul? If so, why is the change required, and how would the recent advancements such as AIEd, ARLE, and AR possibly aid in digital revamping of dental education curriculum?

Does dental education and research require a curriculum overhaul? It is an affirmative 'yes', it does require an overhaul, as there are many caveats and drawbacks that the conventional dental teaching model experiences, and some of the inadequacies are enumerated:

(1) lack of experts or expert's tutorial time,
(2) absence of dental education amalgamation into non-dental areas,
(3) non-assimilation of technological advancements into dental education and research,
(4) scarcity of real world cases to enhance manual dexterity and cognitive comprehension (Roy et al. 2017; Iacopino 2007; Towers et al. 2019; Gandedkar et al. 2021).

Refining skills using conventional dental simulators (e.g., phantom head, plastic teeth), in a model-based preclinical set-up has a long history of not only providing exemplary support in honing psychomotor skills, but also, have been excellent in equipping the dental student for real-life clinical applications (Perry et al. 2015). However, conventional dental simulators pose several shortcomings, such as:

(1) breach in infection control in the form of possible water line contamination (for e.g. Legionella),
(2) lack of plastic teeth's ability to mimic natural teeth and periodontium morphology, composition, and other typical characteristics, in entirety and hence, inadequate sensory feedback,
(3) irreversible, environmentally unfriendly procedures,
(4) non-real-time feedback and evaluation by supervisors (Li et al. 2021).

To offset shortcomings of the conventional dental simulators and to revamp dental education, a plethora of computer-supported and virtual reality (VR) simulators were introduced, in the late 1990s and in early 2000s. VR simulators brought a considerable transformation, in terms of conveying information and empowering gaining of skills, especially in dental students and marked a new era of digital-enabled dental education (Llena et al. 2018).

9.2 Application of VR and AR in Dental Education and Research

The scope of this chapter is to explore the current and future potential of AR, and especially ARs ability to influence 'Dental Education', hence, the applications of dental simulators in general dental fields is analysed sparingly. Several virtual reality systems have been adopted with engagement of wide-ranging sensory channels by virtual simulation, these can be broadly classified into three categories based on the degree in which the range of sensory channels are engaged by virtual simulation, and are:

(1) Fully immersive VR (e.g., head-mounted displays, HMD's),
(2) Semi-immersive VR (e.g., cave automatic virtual environment, CAVE),
(3) Non-immersive screen-based VR (Computer-aided design systems, CAD) (Checa and Bustillo 2020; Gandedkar et al. 2021).

The introduction of VR in dental education has enabled the user not only to wholly engage in a virtual environment, but also to interact with it using a specialized controller device. Monterubbianesi R et al. enumerated some of the most commonly used VR dental simulators, such as, PerioSim®, Dentsim™, Individual Dental Education Assistant (IDEA), Simodont®, Voxel Man, and computerized dental simulator (CDS) (Monterubbianesi et al. 2022). See Table 9.1. Further, the table explains the central parameters such as, type of teeth used (animated or plastic), realism of tactile sensation, ergonomic postures, seamless data transfer of operating procedures between learner and program instructor for facilitation of procedure review, training, grading, and verifying. For example, regarding dental education, VR-based haptic systems such as Individual Dental Education Assistant (IDEA) is constructed to empower students to be adaptable and competent in the use of dental handpieces by performing removal of predesigned virtual materials with different shapes, aiming students to develop manual dexterity (hand flexibility) whilst maintaining speed (Ben-Gal et al. 2013). Similarly, Dental simulators enabled by VR have found many applications in dentistry, for instance:

(1) Endodontics—dental caries detection, caries removal, light-curing skills, and

Table 9.1 Comparison of different dental simulators (reproduced under creative commons, Monterubbianesi et al. 2022)

	PerioSim®	Dentsim™	IDEA[a]	Simodont®	Voxel man	CDS
Teeth used	Animated	Plastic teeth	Animated	Animated	Animated	Animated
Right and left operation	Available	Available	Available	Available	Available	Available
Reported real life experience	Tactile sensation is realistic for teeth but not for gingiva	Realistic experience using plastic teeth on a real manikin	Tactile sensation still needs to be tuned to simulate a genuine sensation	3D images are realistic. However, the texture of healthy decayed and restored tooth structure still needs improvement	/	/
Ergonomic postures	No	Yes	No	Yes	No	Yes
Direct transfer of data to program instructor/tutor	Not available	Yes run time control application enables the instructor to control run time grades	Yes the software contains a replay mode upon completion of a specified task, it can be watched in full by the student or the instructor	Yes allows the instructor to watch six simulators live at once and record all preparations for evaluation in order to give feedback later	Not available	Yes operating procedures are recorded and can be reviewed to facilitate in training, grading and verifying
Instant feed back	No	Yes	Yes	Yes	Yes	Yes
Exam simulation	Yes	Yes	No	Yes	Yes	Yes

[a] Individual dental education assistant

endodontic cavity preparation,

(2) Periodontics—periodontal probing, ultrasonic scaling,
(3) Oral and maxillofacial surgery—dental anaesthesia training, dental extraction skills, orthognathic jaw surgery,
(4) Dental radiography—intra-oral X-ray imaging,
(5) Prosthodontic—tooth preparation,
(6) Implantology—implant scheme design,

(7) Orthodontics—treatment planning, bonding of brackets (Li et al. 2021).

Li et al. enumerates several advantages of dental simulators when compared to phantom-based traditional training methods, by overcoming the deficiencies and providing students with more flexible training times, allowing digital objective evaluation, eliminating the risk of treatment, and enhance the safety of patients (Liu et al. 2020; Nassar and Tekian 2020; Raja'a and Farid 2016; Escobar-Castillejos et al. 2016; Quinn et al. 2003; Dixon et al. 2021).

However, the application of VR in dental education has several drawbacks. Various authors have identified these drawbacks, and they are (1) lack of the transmission rate (using 5G technology) of high-definition video data, and reduce the response delay, (2) dearth of refining simulation force feedback by incorporating cutting-edge technology like deep learning (DL), machine learning (ML) and artificial neural networks (ANN), and (3) inefficiency in seamless visual and tactile response in an interactive milieu to increase the immersive capability of dental simulators (Ben-Gal et al. 2013; Wang et al. 2016; Georgescu et al. 2020; Zheng et al. 2016; Vo et al. 2017).

Dzyuba et al. in their systematic review on virtual and augmented reality in dental education divided the AR into three categories based on the performance of AR in comparison to conventional methods of teaching technical dental operative skills (Dzyuba et al. 2022):

(1) AR applications that out-performed conventional methods
(2) AR applications that equalled the performance of conventional methods
(3) AR applications used adjunctively to conventional methods.

AR applications that out-performed conventional methods

The AR systems used in this category were, Moog Simodont dental trainer (Moog B.V, Nieuw-Vennep, Netherlands) (Murbay et al. 2020), Dental Simulator (Campinas, SP 13083765, Brazil) (Mladenovic et al. 2019) Omni haptic devices (SensAble Inc., Woburn, MA, USA) (Suebnukarn et al. 2012), dental simulator AR mode (v1.13—Campinas, Brazil) (Mladenovic et al. 2020).

AR applications that equalled the performance of conventional methods

The AR systems used in this category were, Force feedback arm device (Geomagic Touch X Haptic Device, Geomagic Inc., Morrisville, NC, USA) (Vincent et al. 2020), Two Omni haptic devices (SensAble, Inc, Woburn, MA, USA) (Dwisaptarini et al. 2018), PHANTOM Omni (SensAble Inc., Woburn, MA, USA) (Suebnukarn et al. 2010), and haptic devices (SensAble Inc., Woburn, MA, USA) (Suebnukarn et al. 2011).

AR applications used adjunctively to conventional methods

The AR systems used in this category were, Simodont VR haptic dental simulator (MOOG, Nieuw-Vennep, Netherlands) (Al-Saud et al. 2017), DentSim (Image Navigation, New York, NY) (Kikuchi et al. 2013).

Some of the AR modalities that have found application in dental education are enumerated in Table 9.2. To discuss the AR application in dental education, the further chapter is divided into following domains of dentistry; Operative and Endodontics, Prosthodontics, Oral and Maxillofacial Surgery (OMFS), Orthodontics, Basic Sciences, and Paediatrics.

9.2.1 Operative and Endodontics

Augmented reality in true sense follows the definition of AR as "the addition of computer-generated output, such as images or sound, to a person's view or experience of his or her physical surroundings by means of any of various electronic devices" (Dzyuba et al. 2022). The above definition stands true as far as application of AR in operative dentistry and endodontics is concerned. VR (e.g., Simodont dental trainer) in general, has found its place in the application in virtually training dental students with the aid of virtually enhanced tooth, jaw shapes, and dental armamentarium including a 'handpiece-replica' that provides senses of touch and proprioception (haptic). However, Dzyuba et al. notes, AR takes the upper hand over VR, as AR provides realism closer to real-life-scenarios, in three ways: (1) computer-tracked navigation, on a (2) live dental model capable of reality-augmentation, and (3) executed in real-time delivering explicit feedback over a monitor. The commonly used AR-guided operative and endodontics dentistry applications are enumerated in Table 9.2.

9.2.2 Basic Sciences Teaching in a Dental School Setting

Basic sciences teaching is one of the primary requisites for introducing the dental student to the anatomical structures and other associated edifice in not only sensitising the student, but also, to invoke interest in the subject. Conventionally, a human cadaver is used for the dissemination of knowledge, however, this mode of teaching is labour intensive, would consume significant time to assimilate, organise, and deliver the course material, and incur high cost, pose occupational health and safety (OHS) problem, and so on. Several modern-day 'synthetic cadaver-based' training systems have attempted to replace the conventional 'deceased body cadaver' (Blackburn et al. 2021). Some of commonly used synthetic cadaver's are—SynDaver®, Laerdal® and AirSim® manikins. Blackburn et al. studied the anatomic accuracy of airway training manikins compared with humans and concluded that manikins do not have anatomically correct static dimensions in relation to humans and these inaccuracies may lead

Table 9.2 Augmented Reality in Dental Education with description of hardware and software used

Dental field	Author, year	Augmented reality (AR) used	
		Hardware	Software
Paediatrics • Oral hygiene practise in children • Local anaesthesia practise in children • Distraction analgesia technique for pain management in children	(Mladenovic et al. 2020)	Syringe with saliva samples were collected by Sarstedt Salivette (Salivette, Sarstedt Inc.). with AR scheme	Dental simulator mobile application in AR mode (v1.13—Campinas, Brazil), AR mode powered by Vuforia Engine (PTC, Parametric Technology Corporation)
	(Atzori et al. 2018)	Oculus rift DK2 and CV1 virtual reality goggles	SnowWorld gaming software MSI GT Series GT72 Dominator Pro G-1252 Gaming Laptop 6th Generation Intel Core i7 6700HQ
Operative and Endodontics dentistry • Endodontic access • Root canal access • Cavity preparation • Preparation of gold onlay restoration	(Suebnukarn et al. 2010)	• Dental simulator (Campinas, SP13083765, Brazil) AR mode of dental simulator mobile application in AR mode (v1.13—Campinas, Brazil) • AR mode powered by Vuforia Engine (PTC, Parametric Technology Corporation)	
	(Suebnukarn et al. 2011)	Two haptic devices (SensAble Inc., Woburn, MA, USA)	The virtual reality (VR) simulator was comprised of a 2.8-GHz Pentium 4 PC, with 256 MB RAM and a 13-in computer monitor
	(Suebnukarn et al. 2012)	Two Omni haptic devices (SensAble Inc., Woburn, MA, USA)	Virtual reality (VR) operates on a laptop with 1.6-GHz Intel processor and 2 GB of main memory
	(Llena et al. 2018)	AR cavity models, laboratory model and the 3D image (AR)	Aumentaty Author 1.2 + Aumentaty Viewer (Bienetec, Valencia, Spain) for computer-based AR and Augment app (Augment, Paris, France)

(continued)

Table 9.2 (continued)

Dental field	Author, year	Augmented reality (AR) used	
		Hardware	Software
	(De Boer et al. 2016)	Simodont dental trainer digital Circular polarised glasses (3D Optix, Gent, Belgium)	Multimedia projectors from LG, type HS 101 (resolution 800 9 600)
	(Al-Saud et al. 2017)	Simodont VR haptic dental simulator (MOOG, Nieuw-Vennep, Netherlands)	Courseware software (developed by the Academic Centre for Dentistry Amsterdam (ACTA), Amsterdam, Netherlands)
	(de Boer et al. 2017)	Moog Simodont dental trainer (Nieuw-Vennep, the Netherlands) virtual learning environment (VLE) with FFB force feedback (FFB)	
	(Dwisaptarini et al. 2018)	Two Omni haptic devices (SensAble, Inc, Woburn, MA, USA)	Fluorescence-aided caries excavation (FACE; Sirona, Hanau-Wolfgang, Germany)
	(de Boer et al. 2019)	Simodont dental trainer (Moog, Nieuw-Vennep, the Netherlands)	
	(Vincent et al. 2020)	Haptic simulator (Virteasy; HRV Simulation, Changé, France) 3D screen stereoscopic glasses (Estar America ESG6100; BienestarAmerica, Hillstar, OR, USA) Force feedback arm device (Geomagic Touch X Haptic Device, Geomagic Inc., Morrisville, NC, USA)	

(continued)

Table 9.2 (continued)

Dental field	Author, year	Augmented reality (AR) used	
		Hardware	Software
Prosthodontics • Implant placement • Facial recognition for smile design • Practising porcelain fused to metal crown preparation • Maxillectomy and prosthetic defect rehabilitation	(Kikuchi et al. 2013)	DentSim (Image Navigation, New York, NY)	Virtual reality simulation (VRS) program
	(Elbashti et al. 2020)		Vuforia software and Unity3D software
Orthodontics • Bracket placement guide • Image-based tracking of the teeth • E-learning and augmented reality (AR) platforms	(Lo et al. 2021)	Wireless handpiece intra-oral camera (QOCA® Q-tube Wi-Fi Teeth Scope Pro, Quanta Computer Inc., Taoyuan City, Taiwan	Digital modeling software Ortho Analyzer (3Shape Dental Systems, Copenhagen, Denmark)
	(Aichert et al. 2012)	• (MAR) AR-assisted bracket	
	(Rao et al. 2017)	• Navigation system	
Oral and maxillofacial surgery (OMFS) • Local anaesthesia administration Inferior Alveolar Nerve Block (IANB) • Administration of Anterior Superior Alveolar nerve block (ASA) • Jaw surgery • Planning and performing repair of cleft lips	(Pulijala et al. 2018)	Oculus Rift and Leap Motion devices (Leap Motion, Inc, San Francisco, CA)	Oculus rift development kit (DK2)
	(Mladenovic et al. 2019)	Syringe with Saliva samples were collected by Sarstedt Salivette (Salivette, Sarstedt Inc.) with AR scheme	AR mode (v1.13—Campinas, Brazil) AR mode powered by Vuforia Engine (PTC, Parametric Technology Corporation)
Basic sciences • Objective Structured Clinical Examination (OSCE), "AR based-OSCE" • Pre-clinical training • Cadaver training	(Murbay et al. 2020)	Moog Simodont dental trainer, Moog B.V, Nieuw-Vennep, Netherlands	Simodont Dental Trainer software V.4
	(Zafar and Zachar 2020)	HoloLens (HoloHuman) by Pearson, 3D4 Medical Ltd, Blackrock, Co. Dublin, Ireland	

to imprecise airway device development, negatively affect training and cause overconfidence in users (Blackburn et al. 2021). AR could potentially overcome these drawbacks by utilising AR's augmented mirrored image modality enabling one's own body's internal structures to not only visualise, but also interact in a tangible manner (Dhar et al. 2021). Dhar et al., in their article on AR in medical education and students' experiences and learning outcomes enumerated several simulation-based platforms, such as, HoloHuman, OculAR SIM, and HoloPatient, HoloChemistry can digitally generate three-dimensional representations to be integrated with real environmental stimuli. Zafar and Zachar studied the application of HoloHuman, HoloLens (HoloHuman by Pearson, 3D4 Medical Ltd, Blackrock, Co. Dublin, Ireland) augmented reality as a novel educational tool in dentistry and concluded that AR offers an additional means of dental anatomy training. However, caution must be exercised as there are several limitations with the device (Zafar and Zachar 2020). See Fig. 9.1.

Fig. 9.1 The AR app 'HoloHuman' showing a virtual cadaver placed on a real examination table. The moderator (shown) can interact with the model and user interface using a HoloLens headset. Structures, organs, and systems can be examined individually or in combination and are fully supported by visual narrative and digital dissection tools (figure reproduced under creative commons) (Dhar et al. 2021)

9.2.3 *Oral and Maxillofacial Surgery (OMFS)*

Farronato et al. (2019) conducted a systematic review to ascertain the current state of use of augmented reality in dentistry. They concluded that most of the studies carried out on AR field of application were in the oral and maxillofacial area (OMFS) (21 out of 23). See Fig. 9.2.

They further recognise that AR application in OMFS is predominantly categorised into three specific areas: implantology, maxillofacial surgery and oral surgery. Additionally, AR systems have found great impact in the subspeciality of OMFS such as, cranio-maxillofacial reconstructive surgery (high fidelity haptic feedback enabling analysis, planning, and preoperative testing) (Olsson et al. 2013), cleft lip and plate repair (computer-based surgical simulation system for planning and performing cleft lip repair) (Schendel et al. 2005), reduction of complex maxillofacial fractures (CHAI3D, Computer Haptics and Active Interface used for rendering the haptic feedback) (Zhang et al. 2015), craniofacial syndromes management (e.g., hemifacial microsomia) via precise positioning of an intraoral distractor (Integration of 3D digital graphics of occlusal splint with marker and cone beam computed tomography image of mandible to establish virtual image of AR) (Qu et al. 2015), and training of Le Fort I osteotomies fixation (AR immersive workbench systems with haptic feedback). Regarding OMFS education and AR application is concerned, Pulijala

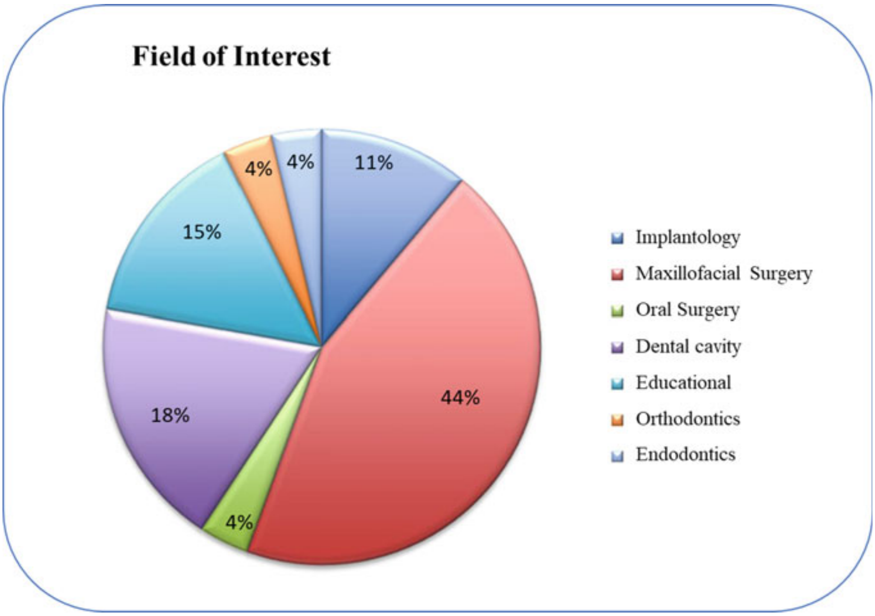


Fig. 9.2 Figure depicting AR/VR applicability in the field of dentistry (figure reproduced under creative commons) (Farronato et al. 2019)

et al. conducted a multisite, randomized controlled trial (RCT) to evaluate the effectiveness of immersive virtual reality in performing the Le Fort I osteotomy while training surgical residents. Immersive VR (iVR) experience was developed using Oculus Rift and Leap Motion devices (Leap Motion, Inc, San Francisco, CA) to address the challenge of novice residents of not having confidence (as high as 40%) in performing a major surgical procedure. The study concluded that iVR experiences improve the knowledge and self-confidence of the surgical residents (Pulijala et al. 2018). Similarly, Mladenovic et al. studied the effectiveness of a mobile augmented reality simulator Dental Simulator (Campinas, SP 13083765, Brazil), for local anaesthesia training with dental students who are administering inferior alveolar nerve block (IANB) and concluded that the students who used the mobile simulator in addition to their education in augmented reality had greater success than the students who used only the conventional educational methods (Mladenovic et al. 2019).

9.2.4 Orthodontics

Rao et al. proposes mobile augmented reality (MAR) to overcome the challenges faced by conventional pedagogical way of imparting orthodontic knowledge in a teaching setting. MAR is an AR platform that is intended to assimilate, disseminate, and retain through three dimensional (3D) real-world visualization and haptic experience. The MAR, according to the authors, not only effectively neutralizes the practical exercise's stressful situation that are typically carried-out in a simulation laboratory, but also motivates the student to learn and encourages deeper learning (Rao et al. 2017, 2020; Iacopino 2007; Ifenthaler and Yau 2019). To achieve the aforesaid, MAR typically integrates three main technological advancements; image recognition, interactive controls, and computer graphics (Rao et al. 2020). Aichert et al. and Lo et al. (Aichert et al. 2012) further explore the possibilities of integrating image-based tracking for precise guided-bracket placement for orthodontic tooth movement (Aichert et al. 2012; Lo et al. 2021). The AR-assisted bracket navigation system comprised two technical modules, namely the facial axis of the clinical crown (FACC) detection module and the bracket bonding navigation module. Pre-treatment digital models prepared from intraoral scanner (TRIOS 3 Basic, 3Shape Dental Systems, Copenhagen, Denmark) were used for virtual setup using the digital modelling software Ortho Analyzer (3Shape Dental Systems, Copenhagen, Denmark) and this digital modelling set-up was used to predict the final setup using AR-assisted bracket navigation system. The study concluded that AR-assisted bracket navigation system improved the accuracy of bracket placement and decreased the procedure time of lab stage (Lo et al. 2021). See Fig. 9.3.

Several other areas of orthodontics have been explored such as Dental monitoring, augmented 3D image visualization, 3D printing utilising VR and AR and so on (Gandedkar et al. 2021; Hansa et al. 2020, 2021). Hansa et al. conducted a study to evaluate the effects of Invisalign clear aligner treatment with and without Dental Monitoring (DM) for treatment duration, number of appointments, refinements and

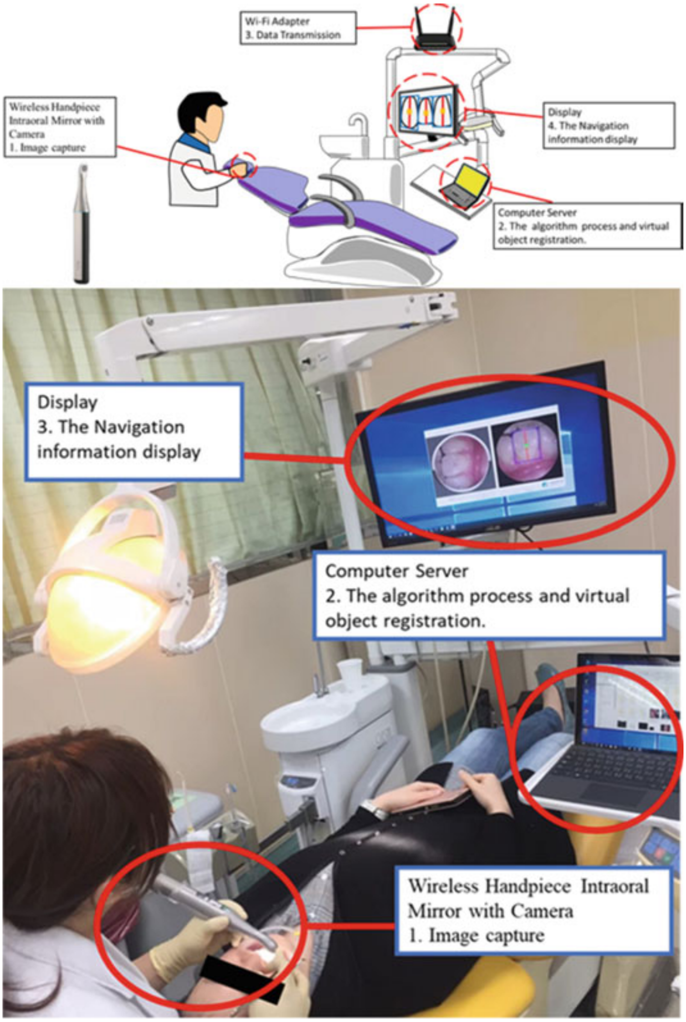


Fig. 9.3 The setting and actual clinical operation of the AR-assisted bracket navigation system (Lo et al. 2021) (reproduced under creative commons)

refinement aligners, and accuracy of Invisalign in achieving predicted tooth positions (aligner tracking). See Fig. 9.4. They concluded that DM group improved aligner tracking in comparison to without DM (Hansa et al. 2020). Also, AR could be used to determine the degree to which this tool accurately reveals anatomical structures and determine whether application of this modality enhances understanding of our outcomes and is efficient for both educating and patient management as far as orthodontics is concerned (Figs. 9.5 and 9.6) (Chng et al. 2019; Gandedkar et al. 2017, 2018, 2021; Vandekar et al. 2015).

Fig. 9.4 Dental monitoring application as a form of advanced remote monitoring solution to build a virtual practice

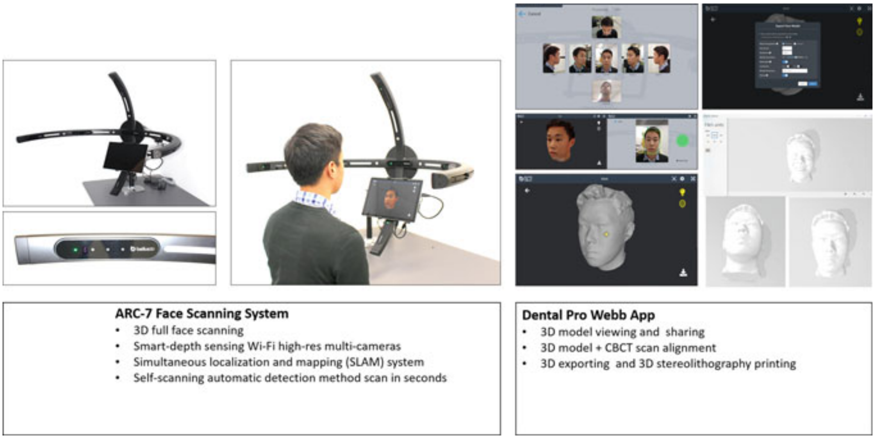


Fig. 9.5 Integration of 3D full face scanning for simultaneous localization and mapping system (SLAM) which further transforms 3D images into 3D stereolithographic printing

9.2.5 Prosthodontics

Researchers from Tokyo Medical and Dental University wanted to understand the purpose of virtual reality simulation (VRS) as a modern-day dentistry education and

Fig. 9.6 Improved visualization of intraoral structures with a possibility of three-dimensional printed models used for the 3D printing of orthodontic appliances

research modality, hence they evaluated its use in understanding different educational programs, particularly, students honing skills related to porcelain fused to metal (PFM) crown preparation. DentSim (Image Navigation, New York, NY), a real-time preclinical simulator with the use of three-dimensional (3D) graphics as a VRS unit was combined with a dentoform. Further, it was integrated with seven tracking light emitting diodes (LEDs), and a standard jaw with standard turbine 650B LUX-3 (KaVo Dental GmbH, Germany) was used. The study indicated that the application of the VRS system increased student training for PFM (Kikuchi et al. 2013). The above is an example of several VRS systems that have been implemented in the education and research (Esser et al. 2006; Welka et al. 2006; Yasukawa 2009; Rees et al. 2007).

The scores gathered for occlusal reduction and wall incline for DentSim groups, in comparison to non-DentSim group, indicated that the students that used VRS needed less supervision, and fewer instructions than it would be needed for the training process and feedback evaluations. Further, the studies elucidated that training students VRS could be used to review self-preparation from various perspective, such as accuracy of their work, preparation times, clinical experience, development of better skills, differentiating characteristics between enamel and dentin during preparation, visually understand the anatomical shape of the structure, and other quantitative and qualitative evaluation that would be beneficial for the erudition and empowerment of the student (Kikuchi et al. 2013).

9.2.6 Paediatrics

Similarly, as previous application of AR, the impact of AR simulator on the awareness of acute stress level in students administering local anaesthesia (anterior superior alveolar nerve (IASAN) to paediatric patients was assessed. The infiltrative anaesthesia solution is deposited via AR mode (v1.13—Campinas, Brazil), available for iOS (App Store) and Android (Google Play Store). AR mode powered by Vuforia Engine (PTC, Parametric Technology Corporation) creates immersive experience using mobile camera or VR glasses. The study concluded that although, augmented reality application improves students administering LA to paediatric patients for the first time, it cannot reduce acute stress. However, AR mobile simulator application can pose as an important advantage to students that transition from the pre-clinical level to clinical skills. AR system is well suited for these factors and that it can be used without many restrictions (Mladenovic et al. 2020) Also, Oculus Rift VR helmet, Oculus Rift DK2 and CV1 virtual reality goggles have also been used (Atzori et al. 2018).

Several immersive interactive ludic-educational interfaces have been tested to motivate oral hygiene practice in children. 3D frameworks, such as Kinect for Windows SDK 2.0 (Microsoft Corp), that visualize and interact in a virtual milieu by providing realistic situation with an idea of a distinct possibility of converting ‘digital naïve’ to ‘digital natives’ whilst achieving the objective (Amantini et al. 2020; Atzori et al. 2018).

9.2.7 Implementation of AR in Dental Education

Although AR application is finding its foothold in the dental arena, however, formal introduction via dental education and training is lacking. In order to properly implement the AR-related dental teaching, Gandedkar et al. (2021) discussed the exploration of a ‘preliminary model’ with (1) scrupulous representation of Virtual Reality (VR), Augmented Reality (AR) and Artificial Intelligence (AI) in education and research of dentistry via; intelligent tutoring systems (ITS), laboratory and preclinical teaching, patient-centred interactive teaching, and research-oriented teaching (Fig. 9.7), and (2) implementation of a synchronous approach of strategic frameworks, professionalism and patient care delivery in the era of augmented reality learning experience (ARLE) interactive multimedia technology (both guided and self-guided) and cognitive and sensorial assimilation of rendered world. See Figs. 9.8 and 9.9.

According to Towers et al., the following quantitative factors must be taken into consideration; (1) feasibility of hardware equipment simulation, (2) the simulation pragmatism and ability to reproduce life-like environ, (3) seamless feedback and evaluation system, and (4) authentication of student-assignments by active role of tutors or educators. Further, especially for the realization of active role of tutors in the AR milieu, one must consider ‘educating the educator’ to help the educator orientate themselves, for the successful implementation of AR pathway in dental education.

Figure 9.10 depicts a blue-print of how one can achieve this by pooling technical and educational designers input as a ‘starting-point’ to identify AR learning needs and objectives. Further, the conceptual prototype or learning material specifically

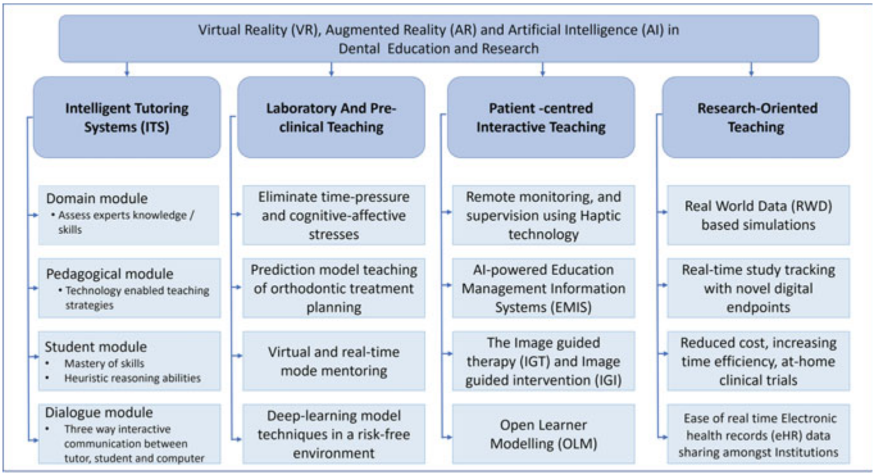


Fig. 9.7 Overview of Virtual Reality (VR), Augmented Reality (AR) and Artificial Intelligence (AI) in dental education and research. "Reprinted from Publication title—Seminars in Orthodontics, 27/2, Narayan H. Gandedkar, Matthew T. Wong, M. Ali Darendeliler, titled Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: An insight, Page No. 9, Copyright (2021), with permission from Elsevier" (Gandedkar et al. 2021)

developed for the dental education is predicted whilst eliciting appropriate subject-specific response and feedback of AR learning from teachers, so that the application of AR experience in dental curriculum is thoroughly appraised.

9.3 Conclusion and Future Forecast

In this chapter, an overview of application of augmented reality, virtual reality and artificial intelligence in dental education is provided with emphasis on features, principal devices, and application of AR. Further, implementation of AR in dental education is discussed in terms of assimilation of both students and tutors' biofeedback (emotional and intelligence quotients along with multi-sensorial attributes) to generate an augmented reality learning experience (ARLE). Monterubbianesi et al. conducted a narrative review on exiting platforms and future challenges regarding Augmented, Virtual and Mixed Reality (MR) in Dentistry (Monterubbianesi et al. 2022). The review concluded that AR can improve clinical practice, enhance inter-professional communication, and offer diagnoses in a better manner. However, the review also cautioned that although modern day technological marvels such as AR, AI, and VR may hold the potential to bring the dentistry education and research to twenty-first century, however, one must understand that their actual potential is yet to be fully realised. Further, to assimilate the AR, VR, and MR into a composite of

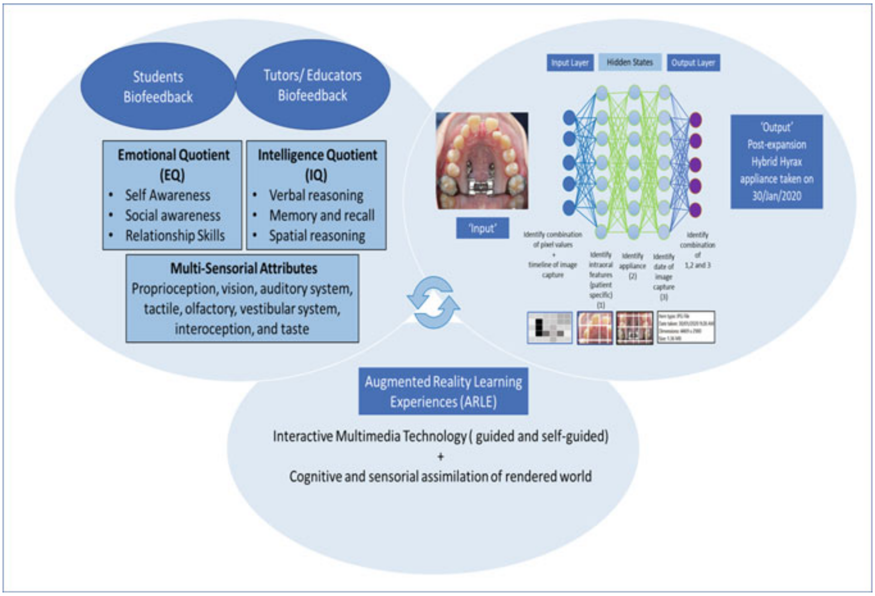


Fig. 9.8 Orthodontic application of artificial neural networks (ANN) multilayers—input, multiple hidden and output layers. For e.g., an input layer of just an intraoral image can identify, via multiple hidden layers with algorithmic combination of pixel values and timeline of image capture, patient's orthodontic treatment stage, mechanics employed with complete description of appliance, morphological characteristics of teeth (along with angulation, inclination, etc.), and date of image capture. Further, the ANNs output information is assimilated with both students and tutors' biofeedback (Emotional and intelligence quotients along with multi-sensorial attributes) to generate an augmented reality learning experience (ARLE). "Reprinted from Publication title—Seminars in Orthodontics, 27/2, Narayan H. Gandedkar, Matthew T. Wong, M. Ali Darendeliler, titled Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: An insight, Page No. 9, Copyright (2021), with permission from Elsevier" (Gandedkar et al. 2021)

extended reality (XR) a careful consideration of ethical aspects, erudition philosophies and intellectual challenges in integrating, and executing in a tripartite environment of 'student-computer-tutor' milieu should be further investigated to improve understanding. See Fig. 9.11.

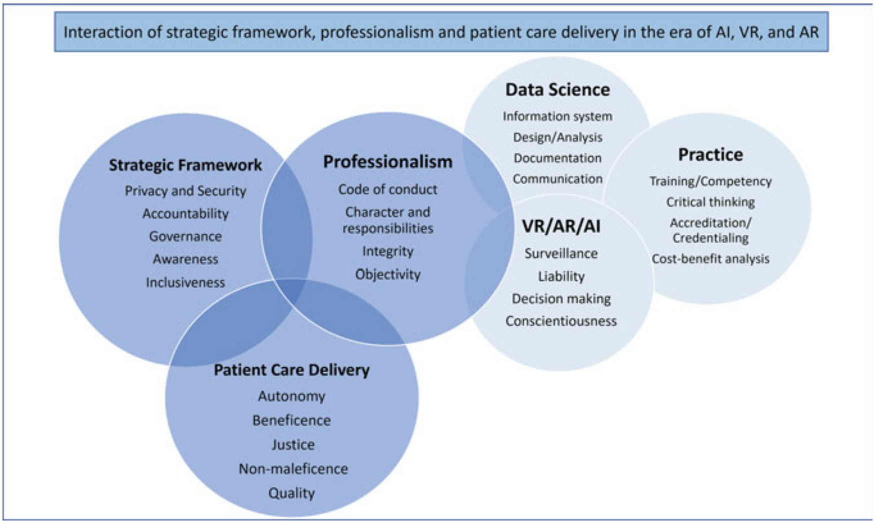


Fig. 9.9 Interaction of strategic framework, professionalism, and patient care delivery in the era of AI, VR, and AR. "Reprinted from Publication title—Seminars in Orthodontics, 27/2, Narayan H. Gandedkar, Matthew T. Wong, M. Ali Darendeliler, titled Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: An insight, Page No. 9, Copyright (2021), with permission from Elsevier" (Gandedkar et al. 2021)

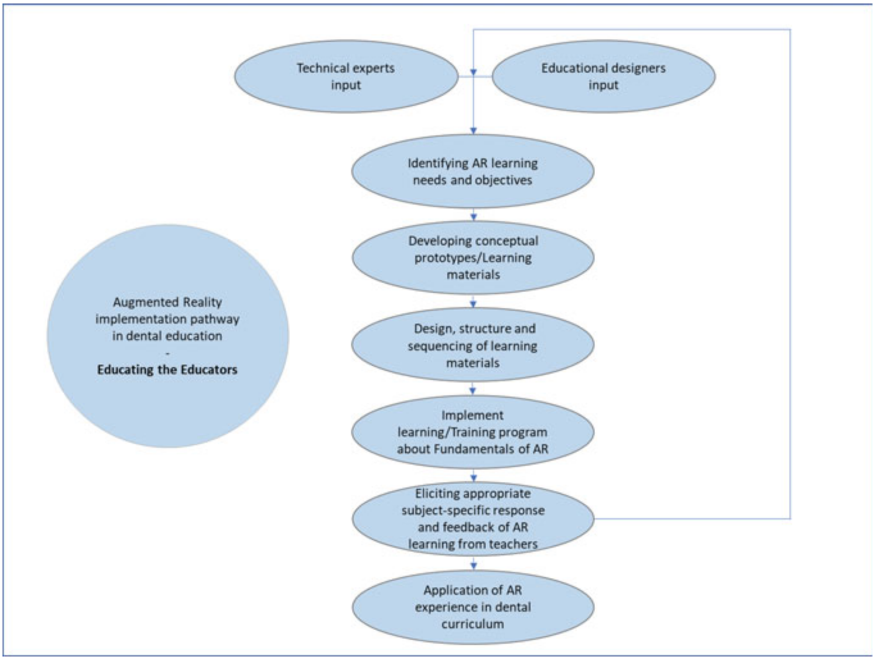


Fig. 9.10 Implementation of AR in dental education—educating the educator

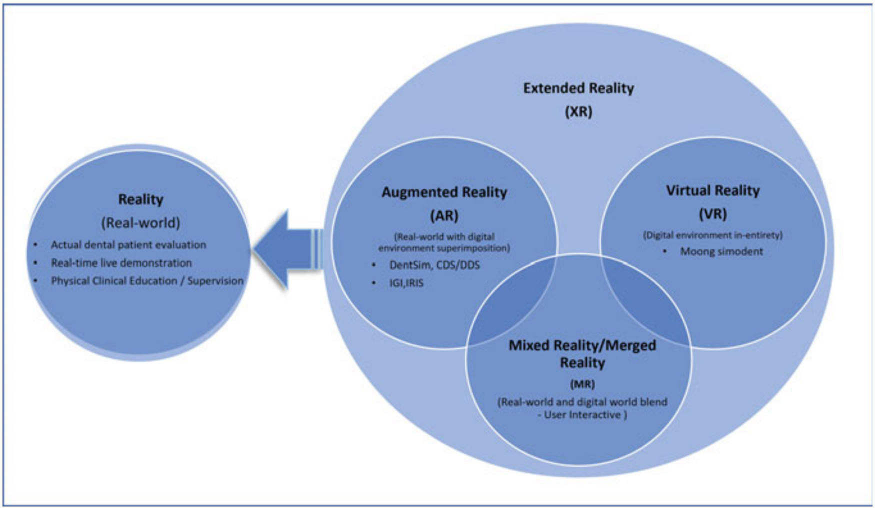


Fig. 9.11 Integration of AR, VR, and MR into a composite of extended reality (XR) to synergise these technological advancements into the real-world as close as possible

References

Aichert A, Wein W, Ladikos A, Reichl T, Navab N (2012) Image-based tracking of the teeth for orthodontic augmented reality. In: International conference on medical image computing and computer-assisted intervention. Springer, pp 601–608

Al-Saud LM, Mushtaq F, Allsop MJ, Culmer PC, Mirghani I, Yates E, Keeling A, Mon-Williams MA, Manogue M (2017) Feedback and motor skill acquisition using a haptic dental simulator. Eur J Dent Educ 21:240–247

Amantini SN, Ribeiro S, Montilha AAP, Antonelli BC, Leite KTM, Rios D, Cruvinel T, Neto NL, Oliveira TM, Machado MAAM (2020) Using augmented reality to motivate oral hygiene practice in children: protocol for the development of a serious game. JMIR Res Protoc 9:e10987

Atzori B, Grotto RL, Giugni A, Calabrò M, Alhalabi W, Hoffman HG (2018) Virtual reality analgesia for pediatric dental patients. Front Psychol 9:2265

Ben-Gal G, Weiss EI, Gafni N, Ziv A (2013) Testing manual dexterity using a virtual reality simulator: reliability and validity. Eur J Dent Educ 17:138–142

Blackburn MB, Wang SC, Ross BE, Holcombe SA, Kempski KM, Blackburn AN, DeLorenzo RA, Ryan KL (2021) Anatomic accuracy of airway training manikins compared with humans. Anaesthesia 76:366–372

Checa D, Bustillo A (2020) A review of immersive virtual reality serious games to enhance learning and training. Multimed Tools Appl 79:5501–5527

Chng CK, Gandedkar NH, Liou EJW (2019) Future of surgery-first orthognathic approach. In: Surgery-first orthodontic management. Springer

de Boer IR, Lagerweij MD, de Vries MW, Wesselink PR, Vervoorn JM (2017) The effect of force feedback in a virtual learning environment on the performance and satisfaction of dental students. Simul Healthc 12:83–90

de Boer IR, Lagerweij MD, Wesselink PR, Vervoorn JM (2019) The effect of variations in force feedback in a virtual reality environment on the performance and satisfaction of dental students. Simul Healthc 14:169–174

De Boer IR, Wesselink PR, Vervoorn JM (2016) Student performance and appreciation using 3D vs. 2D vision in a virtual learning environment. Eur J Dent Educ 20:142–147
Dhar P, Rocks T, Samarasinghe RM, Stephenson G, Smith C (2021) Augmented reality in medical education: students' experiences and learning outcomes. Med Educ Online 26:1953953
Dixon J, Towers A, Martin N, Field J (2021) Re-defining the virtual reality dental simulator: demonstrating concurrent validity of clinically relevant assessment and feedback. Eur J Dent Educ 25:108–116
Drljević N, Wong LH, Botički I (2017) 'Where does my Augmented Reality Learning Experience (ARLE) belong? A student and teacher perspective to positioning ARLEs. IEEE Trans Learn Technol 10:419–435
Dunleavy M, Dede C, Mitchell R (2009) Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. J Sci Educ Technol 18:7–22
Dwisaptarini AP, Suebnukarn S, Rhienmora P, Haddawy P, Koontongkaew S (2018) Effectiveness of the multilayered caries model and visuo-tactile virtual reality simulator for minimally invasive caries removal: a randomized controlled trial. Oper Dent 43:E110–E118
Dzyuba N, Jandu J, Yates J, Kushnerev E (2022) Virtual and augmented reality in dental education: the good, the bad, and the better. Eur J Dent Educ
Elbashti ME, Itamiya T, Aswehlee AM, Sumita YI, Ella B, Naveau A (2020) Augmented reality for interactive visualization of 3D maxillofacial prosthetic data. Int J Prosthodont 33:680–683
Escobar-Castillejos D, Noguez J, Neri L, Magana A, Benes B (2016) A review of simulators with haptic devices for medical training. J Med Syst 40:104
Esser C, Kerschbaum T, Winkelmann V, Krage T, Faber F-J (2006) A comparison of the visual and technical assessment of preparations made by dental students. Eur J Dent Educ 10:157–161
Farronato M, Maspero C, Lanteri V, Fama A, Ferrati F, Pettenuzzo A, Farronato D (2019) Current state of the art in the use of augmented reality in dentistry: a systematic review of the literature. BMC Oral Health 19:1–15
Gandedkar NH, Koo CS, Chng CK, Por YC, Yeow VKL, Sng K-E (2018) Role of the "craniofacial orthodontist" in a "craniofacial team." J Indian Orthod Soc 52:4
Gandedkar NH, Shen KC, Chng CK (2017) Recent advances in orthognathic surgery diagnosis and management: 3D image acquisition, virtual surgical planning, rapid prototyping, and seamless surgical navigation
Gandedkar NH, Wong MT, Ali Darendeliler M (2021) Role of Virtual Reality (VR), Augmented Reality (AR) and Artificial Intelligence (AI) in tertiary education and research of orthodontics: an insight. In: Seminars in orthodontics. Elsevier
Georgescu M-I, Ionescu RT, Verga N (2020) Convolutional neural networks with intermediate loss for 3D super-resolution of CT and MRI scans. IEEE Access 8:49112–49124
Geroimenko V (2020) Augmented reality in education: a new technology for teaching and learning. Springer
Hansa I, Katyal V, Ferguson DJ, Vaid N (2021) Outcomes of clear aligner treatment with and without Dental Monitoring: a retrospective cohort study. Am J Orthod Dentofac Orthop 159:453–459
Hansa I, Semaan SJ, Vaid NR (2020) Clinical outcomes and patient perspectives of Dental Monitoring® GoLive® with Invisalign®—a retrospective cohort study. Prog Orthod 21:1–7
Iacopino AM (2007) The influence of "new science" on dental education: current concepts, trends, and models for the future. J Dent Educ 71:450–462
Ifenthaler D, Yau J-K (2019) Higher education stakeholders' views on learning analytics policy recommendations for supporting study success. Int J Learn Anal Artif Intell Educ: IJAI 1:28–42
Kikuchi H, Ikeda M, Araki K (2013) Evaluation of a virtual reality simulation system for porcelain fused to metal crown preparation at Tokyo Medical and Dental University. J Dent Educ 77:782–792
Li Y, Ye H, Ye F, Liu Y, Lv L, Zhang P, Zhang X, Zhou Y (2021) The current situation and future prospects of simulators in dental education. J Med Internet Res 23:e23635

Liu, L, Zhou R, Yuan S, Sun Z, Lu X, Li J, Chu F, Walmsley AD, Yan B, Wang L (2020) Simulation training for ceramic crown preparation in the dental setting using a virtual educational system. Eur J Dent Educ 24:199–206

Llena C, Folguera S, Forner L, Rodríguez-Lozano FJ (2018) Implementation of augmented reality in operative dentistry learning. Eur J Dent Educ 22:e122–e130

Lo Y-C, Chen G-A, Liu Y-C, Chen Y-H, Hsu J-T, Yu J-H (2021) Prototype of augmented reality technology for orthodontic bracket positioning: an in vivo study. Appl Sci 11:2315

Luckin R, Holmes W, Griffiths M, Forcier LB (2016) Intelligence unleashed: an argument for AI in education

Mladenovic R, Dakovic D, Pereira L, Matvijenko V, Mladenovic K (2020) Effect of augmented reality simulation on administration of local anaesthesia in paediatric patients. Eur J Dent Educ 24:507–512

Mladenovic R, Pereira LAP, Mladenovic K, Videnovic N, Bukumiric Z, Mladenovic J (2019) Effectiveness of augmented reality mobile simulator in teaching local anesthesia of inferior alveolar nerve block. J Dent Educ 83:423–428

Monterubbianesi R, Tosco V, Vitiello F, Orilisi G, Fraccastoro F, Putignano A, Orsini G (2022) Augmented, virtual and mixed reality in dentistry: a narrative review on the existing platforms and future challenges. Appl Sci 12:877

Murbay S, Chang JWW, Yeung S, Neelakantan P (2020) Evaluation of the introduction of a dental virtual simulator on the performance of undergraduate dental students in the pre-clinical operative dentistry course. Eur J Dent Educ 24:5–16

Nassar HM, Tekian A (2020) Computer simulation and virtual reality in undergraduate operative and restorative dental education: a critical review. J Dent Educ 84:812–829

Olsson P, Nysjö F, Hirsch J-M, Carlbom IB (2013) A haptics-assisted cranio-maxillofacial surgery planning system for restoring skeletal anatomy in complex trauma cases. Int J Comput Assist Radiol Surg 8:887–894

Park WJ, Park J-B (2018) History and application of artificial neural networks in dentistry. Eur J Dent 12:594

Perry S, Bridges SM, Burrow MF (2015) A review of the use of simulation in dental education. Simul Healthc 10:31–37

Pulijala Y, Ma M, Pears M, Peebles D, Ayoub A (2018) Effectiveness of immersive virtual reality in surgical training—a randomized control trial. J Oral Maxillofac Surg 76:1065–1072

Qu M, Hou Y, Xu Y, Shen C, Zhu M, Xie L, Wang H, Zhang Y, Chai G (2015) Precise positioning of an intraoral distractor using augmented reality in patients with hemifacial microsomia. J Cranio-Maxillof Surg 43:106–112

Quinn F, Keogh P, McDonald A, Hussey D (2003) A study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students. Eur J Dent Educ 7:164–169

Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I (2020) A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 147:103778

Raja'a M, Farid F (2016) Computer-based technologies in dentistry: types and applications. J Dent (Tehran, Iran) 13:215

Rao GKL, Iskandar YHP, Mokhtar N (2020) Enabling training in orthodontics through mobile augmented reality: a novel perspective. In: Teaching, learning, and leading with computer simulations (IGI Global)

Rao GKL, Mokhtar NB, Iskandar YHP (2017) An integration of augmented reality technology for orthodontic education: case of bracket positioning. In: 2017 IEEE conference on e-learning, e-management and e-services (IC3e). IEEE, pp 7–11

Rees JS, Jenkins SM, James T, Dummer PM, Bryant S, Hayes SJ, Oliver S, Stone D, Fenton C (2007) An initial evaluation of virtual reality simulation in teaching pre-clinical operative dentistry in a UK setting. Eur J Prosthodont Restor Dent 15:89–92

Renz A, Krishnaraja S, Gronau E. Demystification of artificial intelligence in education

Roy E, Bakr MM, George R (2017) The need for virtual reality simulators in dental education: a review. Saudi Dent J 29:41–47

Schendel S, Montgomery K, Sorokin A, Lionetti G (2005) A surgical simulator for planning and performing repair of cleft lips. J Cranio-Maxillof Surg 33:223–228

Suebnukarn S, Haddawy P, Rhienmora P, Jittimanee P, Viratket P (2010) Augmented kinematic feedback from haptic virtual reality for dental skill acquisition. J Dent Educ 74:1357–1366

Suebnukarn S, Hataidechadusadee R, Suwannasri N, Suprasert N, Rhienmora P, Haddawy P (2011) Access cavity preparation training using haptic virtual reality and microcomputed tomography tooth models. Int Endod J 44:983–989

Suebnukarn S, Rhienmora P, Haddawy P (2012) The use of cone-beam computed tomography and virtual reality simulation for pre-surgical practice in endodontic microsurgery. Int Endod J 45:627–632

Towers A, Field J, Stokes C, Maddock S, Martin N (2019) A scoping review of the use and application of virtual reality in pre-clinical dental education. Br Dent J 226:358–366

Vandekar M, Fadia D, Vaid NR, Doshi V (2015) Rapid prototyping as an adjunct for autotransplantation of impacted teeth in the esthetic zone. J Clin Orthod: JCO 49:711–715

Vincent M, Joseph D, Amory C, Paoli N, Ambrosini P, Mortier É, Tran N (2020) Contribution of haptic simulation to analogic training environment in restorative dentistry. J Dent Educ 84:367–376

Vo N-S, Duong TQ, Tuan HD, Kortun A (2017) Optimal video streaming in dense 5G networks with D2D communications. IEEE Access 6:209–223

Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D, Liang D (2016) Accelerating magnetic resonance imaging via deep learning. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI). IEEE, pp 514–517

Welka A, Spliethа Ch, Wierinckc E, Gilpatrickd RO, Meyera G (2006) Computer-assisted learning and simulation systems in dentistry—a challenge to society Computergestützte Lern-und Simulationssysteme in der. Int. J. Comput. Dent. **9**, 253–265

Yasukawa Y (2009) The effectiveness of cavity preparation training using a virtual reality simulation system with or without feedback. Kokubyo Gakkai Zasshi J Stomatol Soc Jpn 76:73–80

Zafar S, Zachar JJ (2020) Evaluation of HoloHuman augmented reality application as a novel educational tool in dentistry. Eur J Dent Educ 24:259–265

Zhang J, Li D, Liu Q, He L, Huang Y, Li P (2015) Virtual surgical system in reduction of maxillary fracture. In: 2015 IEEE international conference on digital signal processing (DSP). IEEE, pp 1102–1105

Zheng H, Fang L, Ji M, Strese M, Özer Y, Steinbach E (2016) Deep learning for surface material classification using haptic and visual information. IEEE Trans Multimed 18:2407–2416

Chapter 10
The History of Furniture Objects: An Intelligent Augmented Reality Application

Livia Ştefan, Dragoş Gheorghiu, Marius Hodea, and Mihaela Moţăianu

Abstract This chapter investigates how AR processes and user experience can be improved with Machine Learning (ML) techniques and proposes a solution for an intelligent AR mobile application named "FURNITURE". The fusion of AI with AR is not a new initiative but at the same time, its potential is yet to be maximized. Currently, Artificial Intelligence (AI) is applied to almost every domain. As is the case with other technologies, Augmented Reality (AR) can take advantage of the current advancements in AI to allow the creation of a new class of AR applications. The FURNITURE application was designed with two complementary functionalities: (a) to use an ML trained model to decide if an image of a piece of furniture corresponds to a pre-defined style, i.e. is authentic; (b) to augment the piece of the furniture shown in a real context with a history of that furniture style and a 3D model representative for that furniture style, the augmentation being conditioned by a previous validation of authenticity. The "FURNITURE" application uses Google AutoML framework to train an ML model for image classification, and Wikitude AR framework for the development of the mobile AR application. The designed purpose of the AR application is to support students in the fields of art, interior design, architecture and art history, professionals, and the public. The application demonstrates the concept and is open for future extensions, by expanding the object list that can be classified as corresponding to a certain style.

L. Ştefan
Bucharest, Romania

D. Gheorghiu (✉)
Doctoral School, National University of Arts, Bucharest, Romania
e-mail: dragos.gheorghiu@unarte.org

Instituto Terra e Memória, Mação, Portugal

M. Hodea
University of Theatre and Film, Bucharest, Romania

M. Moţăianu
National University of Arts, Bucharest, Romania

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_10

10.1 Introduction

Artificial Intelligence (AI) is a computer science domain related to modelling human intelligent processes such as reasoning, decision-making, visual perception, speech, or handwriting recognition. According to (Encyclopedia Britannica/AI 2022), Artificial Intelligence is "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings". The term is usually associated with systems "endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience" (Encyclopedia Britannica/AI 2022). AI "is not just about robots" but is concerned with "understanding the nature of intelligent thought and action using computers" (Barceló 2009).

Currently, AI is applied in almost every domain, such as industry, medicine, business archaeology, education, or social media, to solve different problems, such as managing complex or tedious processes, assisting users or interpreting information from large data volumes (Collins et al. 2021). For example, in education, AI is "enhancing tools and instruments" (Popenici and Kerr 2017), or can contribute to knowledge construction, using "multichannel data (voice, image, speech, text) and new sensor technologies" (Niemi 2021).

Machine learning (ML) is an important branch of AI that leverages data and algorithms to support "intelligent" processes or behaviours. In archaeology and cultural heritage, ML is used to perform visual analysis (Barceló 2014), assisted modelling (Konečný et al. 2016), or recognition and classification of historic artefacts (Gualandi et al. 2021). Search engines, recommendation engines or conversational chatbots (Google Assistant 2022) incorporate ML techniques and are integrated into numerous everyday applications.

Augmented Reality (AR) technology is based on image recognition and visual analysis in real time. These can be considered as some intelligent processes. More precisely, Computer Vision as a domain of AI used by AR is based on ML algorithms "that enable computers to derive meaningful information from digital images and videos" (IBM CV 2022). Along with the evolution of AR technology, other processes with an advanced degree of intelligence have been integrated, such as visual anchors, depth understanding, environmental recognition (Google AR Fundamentals 2022) or multiple trackers (Wikitude Trackers 2022). All these processes consist in the analysis of shapes, colours, ambient lighting, but they do not achieve the meaningful knowledge of the image content through inference or reasoning, which are AI/ML specific mechanisms.

With the collaboration between AR and AI/ML techniques, a reciprocal improvement of both technologies can be achieved, in both directions: (a) AI/ML to support complex AR processes, e.g., object recognition, environmental understanding, or to create new user experiences (e.g., gesture or voice recognition); (b) AR, with the provision of real-time video feed, can deliver input data to existing ML models, contributing to their continuous and dynamic improvement. The fusion of AI with AR is not a new initiative (Barceló 2006; Bogdanovych et al. 2016) but, at the same time, its potential is yet to be optimized.

This chapter investigates how AR processes and user experiences can be improved by leveraging ML techniques and proposes a solution for developing an “intelligent” mobile AR (MAR) application. The solution was designed starting from the need to develop a history of the furniture for interior design or art history teachers working with large groups of students, as well as for professionals or any user present on site in a museum, antique shop or in another similar context. With the proposed mobile application, any image of a furniture object will be checked as authentic against a selected furniture style, with the help of ML image classification algorithms. The AR augmentations would then be triggered, following the validation of the object’s authenticity.

For the implementation of the application that fuses AI/ML with AR, current AR and AI frameworks and platforms compatible with mobile devices were researched. Despite the evolution of both technologies, the combined use of AR and AI on mobile devices is still a work in progress and is accessible only to those with advanced processing capabilities. However, the performance of mobile applications can be supplemented with external resources, e.g., those offered by cloud platforms. In the case of the current research, ML libraries from Google Cloud Platform (GCP) were considered to train an ML model. The AR mobile application was made with the Wikitude AR framework.

The chapter is organized as follows: a section related to the purpose of the current research work; a section containing a survey of how different technologies are currently fused and how this is impacting the applications and technologies themselves; a section summarizing the current state of AR technology; a section related to a short review of most important concepts in ML and the available ML tools from Google; a section dedicated to the detailed presentation of the research work, having, as a result, a mobile application named “FURNITURE”, and final conclusions.

10.2 An Application for the Design History of Furniture

Throughout history, the fundamental functions of furniture have remained almost unchanged, only the materials and the styling of the forms have undergone changes according to cultural changes.

If the recent history of interior design around the world is analysed (Foyr 2022), it can be observed a tendency to introduce into the present (in interior architecture/interior design) the values of the past, in order to create an eclectic style based on the value of tradition. Interior designers and collectors, such as Axel Vervoordt (2022), show us the augmentative beauty of objects from the past brought into a modern environment.

This requires a good knowledge of the history of the objects used in the setting, respectively of the furniture from the recent or more distant past. The history of habitat planning objects requires a complex presentation, which includes not only the description of the object (Furniture, Styylish History 2022), but also of the cultural and historical environment and not least of its acceptance in the present, in the form of authentic objects or contemporary copies. A quick internet search can provide useful data on “furniture design trends through the ages” such as the “Design + Deliver” blog (Brosa 2022) or Bassett (2022), but not enough to understand the complexity of the shape and history of the object.

Currently, there are 3D furniture databases, which allow a good understanding of the shape of the furniture pieces without providing explanatory data, and which do not contain many pieces of old (e.g. from the sixteenth to eighteenth centuries) furniture (see, for example: 3dmdb 2022; Grabcad 2022; Sketchfab Antique 2022; Open3dmodel 2022; Cgtrader 2022; Designconnected 2022).

In this sense, currently in the digital world, this type of documentation is relatively incomplete. The first action that can be taken is to create new databases and applications that contain both the three-dimensional shapes and explanatory data about the furniture objects, which can be accessed during specialized courses or simply by researching by non-experts. This would allow a process of evolution, in which the basic information taught in the course can be developed by students outside of class, with the help of a dedicated application that the authors propose as part of the research carried out. This application would also allow educating the public both from an aesthetic and historical point of view, as well as for an informed choice of furniture pieces. And last but not least, the application can be a working tool for professionals from different cultural sectors, in the conditions created by the contemporary society where quality, reliable information may be required on short notice.

All these premises can be a challenge for a typical AR application, that can be tackled by extending the AR application with cloud-computing techniques for AI/ML functions and databases. This fusion of AR with AI/ML is addressed in the current chapter through the design of an intelligent application.

10.3 Fusion of Technologies

Since 2017, a Gartner study classifies AI/ML technologies in the category of applied intelligence in a list of top 10 future technologies, while AR/VR is classified in the broader category of “digital” applications (Gartner Report 2017). Another category identified by the Gartner study is represented by the combination of technologies (“mesh”), which gives rise to complex digital platforms and services. Therefore, it can be said that the fusion of technologies with reciprocal enhancements is not a new concept.

10.3.1 Fusion of Technologies in Augmented Reality

In the case of AR, the combination with VR has led to new user experiences and applications, generically called Mixed Reality (MR) (Azuma et al. 2001) or Extended Reality (XR) (Rauschnabel et al. 2022) to refer to immersive technologies that allow a mix between the physical and virtual worlds.

Mobile AR (MAR) uses numerous technologies provided by mobile devices that are fused to produce the AR experience, such as GPS sensors, digital compasses, front and rear video cameras, accelerometers, and gyroscopes. The latest devices are equipped with depth-sensing hardware, such as light detection and ranging (LiDAR) to improve precision in environment analysis and object detection.

Location-based AR was limited to outdoor applications when only the GPS sensor was present on mobile devices. With the modern technologies such as Visual Positioning System (VPS), Bluetooth Low Energy (BLE) (Bluetooth BLE 2022), Low-range Wi-Fi, Ultra-WideBand (UWB) (Android UWB 2022) it is possible to not only develop location-based indoor applications but also to improve outdoor position localization.

10.3.2 Fusion of AI and AR Technologies

AR is a technology that is based on complex processing with real-time output information. AI through ML techniques can increase the performance of AR processes in speed and accuracy, e.g., environmental understanding or detection and classification of objects in complex real-world scenes.

Motion tracking AR processes can be improved with AI techniques through which the AR process is informed about the tracked image or object. As an example, a trained Google ARCore ML model and Google Cloud Vision API can collaborate to identify real-world objects using the real-time video camera feed and can also classify objects by attaching a label to the object in the virtual scene (Google ML 2022). Furthermore, ML can contribute with new AR experiences through innovative information channels, e.g., voice recognition, text recognition, recognition of gestures or facial expressions (Lee and Wong 2020).

A notorious example of AR fused with AI is the AR game Pokémon GO (Pokemongo 2022), in which imaginary creatures (Pokémons) are displayed on a map in an AR view. AI/ML algorithms are used to help generate and localize the Pokémons, after interpreting the environment characteristics (such as objects or landmarks), improve the aesthetics of the Pokémon characters, and also perform continuous game improvements based on the information provided by users (Pokemongo Success 2022).

10.4 Mobile AR Technology: Current State

A common definition of AR is given by (Apple 2022): "Augmented reality (AR) describes user experiences that add 2D or 3D elements to the live view from a device's camera in a way that makes those elements appear to inhabit the real world". AR was also defined in numerous research works as this literature review has shown (Cipresso et al. 2018). For the development of an AR application, several technologies can be used depending on the mobile device's capabilities, and on the complexity and purpose of the application.

Currently, the following choices are available to choose for developing a mobile AR application that leverages the advancements of the technology and mobile devices: (a) native AR frameworks Google ARCore (Google ARCore SDK 2022) and Apple ARKit (Apple ARKit SDK 2022); (b) cross-platform applications, for example, Vuforia (2022) and Wikitude (2022). ARCore 1.33.0 SDK (Google ARCore SDK 2022) introduced several powerful capabilities, among which advanced scene processing, terrain anchors, cloud anchors, ARCore Geospatial API and the integration with ML SDK based on TensorFlow Lite ML framework.

The latest version of ARKit 6 SDK (Apple ARKit 2022) introduces improvements for advanced scene processing, such as location anchors, depth API, raycasting API, and several tools, such as SceneKit and RealityKit—"a custom-built algorithm with Graph data structure, [...] quality assurance and testing AR features". Depth API utilizes one of the most powerful hardware features for AR that are available on modern mobile devices, i.e., LiDAR scanner, to perform a better analysis of scenes and real-world objects and to display virtual objects with greater accuracy.

Wikitude AR framework (Wikitude 2022) provides proprietary algorithms and also integrates with ARCore and ARKit SDKs and with Unity AR Foundation (Unity 2022). The current version is extending the functionality of ARKit and ARCore with more powerful features (Wikitude ARCore 2022) not present in native AR frameworks, e.g., with the capability of video augmentations. The Wikitude AR framework provides a built-in connection to these native SDKs through a new feature called "AR Bridge": "when enabled, the camera configured as AR camera will be driven by AR Bridge", while "the augmentations will be driven by the Wikitude SDK" (Wikitude Concepts 2022).

Native augmented reality applications require more programming skills but have the advantage of allowing developers to make use of all capabilities of the device. A cross-platform application can be developed faster, with less development effort.

10.5 Intelligent Applications with Machine Learning

A formal and general definition of AI-based systems is given by (Janiesch et al. 2021, p. 685): "the capacity of such systems for advanced problem solving [...] based on analytical models that generate predictions, rules, answers, recommendations".

Machine learning is a component of artificial intelligence that "automatically learn meaningful relationships and patterns from examples and observations" (Bishop 2006), and not from rules (Janiesch et al. 2021). Arthur Samuel from IBM defined machine learning as a "field of study that gives computers the ability to learn without being explicitly programmed" (Samuel 1959). Consequently, ML algorithms that are behind the current intelligent applications, are modelling human intelligence, mainly the way humans learn and reason. On big volumes of data, machine learning is used to perform predictions or to highlight key insights or trends. The algorithms are improved gradually using new data sets, through a process called "model training".

Neural networks (NN) and deep learning are ML algorithms used for learning processes, in order to detect, recognize or classify items or objects. When data sets are labelled with information related to the type of objects, the process is called supervised learning. Deep learning algorithms are an improvement of neural networks, that require big data volumes and no human intervention for the process of classification (Janiesch et al. 2021). Due to the fact that ML is not based on rules, it provides the capability for continuous improvement, by moving from "supervised" learning processes to "self-learning" or unsupervised learning processes (Sarker 2021).

ML deep learning and neural networks are heavily used in computer vision. As an example, in cultural heritage computer vision is employed "to ameliorate to way archaeology can deal with the analysis and explanation of the most usual visual marks: shape and texture" (Barceló 2014), or to recognize and classify historic ceramics (Gualandi et al. 2021).

The authors of this chapter have investigated the following ML tools from Google: (a) MLKit for mobile devices ("on-device ML") (Google MLKit 2022) that can be integrated with Google ARCore SDK; (b) AI/ML tools from the Google Cloud Platform (GCP), i.e., the Vertex platform (Google Vertex 2022).

On-device ML (Google MLKit 2022) is based on a light variant of the well-known ML framework, TensorFlow, that allows machine learning models to be implemented on Android and iOS devices, and other consumer devices. This innovation is supported by the current evolution in computing resources provided by modern mobile devices, dedicated to graphical computing, such as Graphical Processing Unit (GPU) or dedicated ML accelerator blocks.

From the analysis made by the authors it resulted that currently MLKit libraries offer limited capabilities in the field of Computer Vision (Image Detection and Classification), that would necessitate a customization in TensorFlow for the purpose of our research work. As an advantage, Google's on-device ML is an alternative requiring no costs, that allows inference with models directly on a device and a direct integration with ArCore SDK.

Google provides a unified cloud AI platform, Vertex (Google Vertex 2022) and an ML framework called AutoML (Google AutoML 2022), that allows users to create and train a model with minimal technical effort, as well as the possibility of generating ML models that are able to run on mobile devices (Google AutoML Deployment 2022).

The advantage of using cloud libraries is that they are much more advanced, but on the other hand, the cloud computing and storage have additional costs. The AutoML framework (Google Vertex Overview 2022) performs ML model training on the Vertex platform at no cost. For the purpose of the current research, AutoML with a free trial usage of Google cloud services was chosen.

10.6 The "FURNITURE" AR-AI Application

10.6.1 The Concept of the Application

During the research work concerning the fusion of AR and AI technologies, an intelligent mobile application was designed having two complementary functionalities: (a) showing if an image of a piece of furniture corresponds to a pre-defined style, i.e. is authentic, by using an ML trained model; (b) augmenting the piece of the furniture shown in a real context with a history of that furniture style and a 3D model representative for that furniture style, once the authenticity of the furniture piece was validated.

The validation is expressed as a percentage, i.e., a score of similarity. Values between 70 and 100% are considered as satisfactory for validation. In other words, a similarity is recognized based on the essential characteristics or traits of the shape of the furniture, with the non-essential characteristics (e.g., colour) being ignored. The history details can be extended with some commercial information.

In the previous research works (Gheorghiu and Ştefan 2019; Gheorghiu et al. 2020, 2021), objects or images identical to the reference images were scanned and augmented. Also, a limited number of images were used, which was appropriate for the purpose declared in the authors' previous research. This approach is considered "typical" for AR applications that use computer vision.

For the current research, the AR augmentations were designed to be triggered for objects or images similar and not identical to the reference ones, and also conditioned by the result of the ML process. The application was designed to offer a mode of operation familiar to users of mobile applications. Thus, the usage and usefulness of the application could be extended to a diverse audience, made up of several categories of users, such as students, professionals, and the public.

For objects that need to be validated by the ML model as authentic, users can obtain the images from various sources (e.g., magazines, internet), or they can produce themselves by using the camera of the mobile device.

The following processes of an application named "FURNITURE" were designed to achieve the purpose of the current research:

(1) **The AI/ML process**

The AI/ML process is composed of 2 sub-processes: (a) ML model training for simple recognition and classification of images showing furniture objects, i.e., if these belong or not to a selected furniture style. The process needs a data set of at least 100 images. The resulting ML model will be integrated into the "FURNITURE" application to be used. As more images will become available, the ML model will be re-trained to improve image recognition and classification accuracy. For the first stage, the image will be considered validated ("true positive") if the recognition score will be between 70 and 100%; (b) execution of the ML model and the resulting similarity score.

(2) **The AR process**

The AR process will be prepared by selecting a very good image, representative of the selected furniture style, and the AR augmentations that will be attached to it. The particularity is that the AR process will be triggered only in the case of an image validation with at least a 70% similarity score.

(3) **The user experience process**

The user experience in using the "FURNITURE" mobile application is related to the designed functionalities. When the mobile application is launched, the images made or taken by the user will be displayed. The selection of an image will make visible a button for validating the image. If no images exist, a message will be displayed for the user to provide photos, name them "furniture1", "furniture2", etc., and save them in a dedicated folder, for example: /furniture/object_recognition/.

The validation button will call the ML model to be executed with the selected image as a parameter (thus the application will enter the AI/ML process). The image will be classified if it fits into the furniture style for which the ML model was trained. If the validation will give a similarity score of at least 70%, then a green mark will be displayed over the image and a button for entering the AR process will be visible. If the validation result is below 70%, only a red mark will be displayed over the image (the AR process will be disabled). In other words, the AR process will be conditioned by the recognition of the authenticity of the object.

Entering the AR process will determine the opening of the device's video camera, in which the related augmentations will be overlayed over the real context with the furniture object or image of it, consisting in the 3D model and the historical and stylistic explanations summarized under the form:

- the full name and the description of the object;
- its cultural and historical context;
- its physical context;
- the location of the museums or antique shops containing the furniture piece;
- the location of shops selling modern copies of the piece (optional).

10.6.2 A Furniture Object and Its Augmentations

In the case of the Guéridon table (Encyclopædia Britannica/Guéridon 2022) the furniture style selected for the current research, the augmentation regarding the historical and stylistic explanations will display the following information:

(a) The Guéridon is a small table with a round top, supported by straight or curved legs. It can also present anthropomorphic decorations in carved wood or gilt bronze. We do not include here those versions on wheels, used in the medical field, or for serving in restaurants.
(b) It appeared in France in the middle of the seventeenth century and had a great spread, being made of different materials. For example, Louis Vuitton made it from Bakelite and iron in 1930.
(c) Although it was made four centuries ago, this piece of furniture is still present in contemporary settings. In the present case, a Guéridon table from the end of the nineteenth century was used to decorate the interior space of a contemporary office.
(d) An importer of this furniture in Romania is the company La Brocante, Albota.[1]
(e) A local producer of modern-style gueridons is the company Stejarmasiv.[2]

10.6.3 3D Models and Images

To implement the concept of the "FURNITURE" application, a Guéridon table was scanned in 3D. This piece of furniture was produced at the end of the nineteenth century—the beginning of the twentieth century in France and is characterized by a simplicity of decoration. It was chosen because it respects the proportions and style of this type of furniture.

The 3D model for the Guéridon was obtained by using photogrammetry, which consists of a series of techniques for creating a digital replica of a real object from multiple overlapping images taken from different perspectives and angles. The Guéridon table was captured using a turntable, having behind it a wall which acted as a background for a later separation from the object (Fig. 10.1).

The workflow comprised several operations: image pre-processing in Photoshop (Adobe Photoshop 2022), e.g., elimination of the background; image processing in Agisoft Metashape (Agisoft 2022); 3D model generation in Agisoft; image post-processing using different software tools such as Autodesk Meshmixer (Meshmixer 2022), Instant Meshes (Instant Meshes 2022) and Agisoft Metashape, the latest also being also used to render the texture. The final 3D model was exported to Sketchfab (Sketchfab 2022), an online platform providing a 3D model database (Fig. 10.2).

[1] http://labrocantearta.ro/?page_id=3480.
[2] www.stejarmasiv.ro/gueridon/.

Fig. 10.1 The Guéridon table prepared for photogrammetry (photo by M. Hodea)

10.6.4 The Design of the Application

The design of the "FURNITURE" application was made in such a way as to allow the user an intuitive experience of going through the ML and AR processes. Below are described the screens of the application, related to the designed functionalities.

(1) **Photographing the furniture object in context**

Object similar to the reference (Fig. 10.3)

Object different from the reference (Fig. 10.4)

(2) **Initiation of the ML process for a selected furniture image**

By clicking the "Validate" button, the classification and validation of the image using the ML model will be initiated.

Object similar with the reference (Fig. 10.5)

Object different from the reference (Fig. 10.6)

(3) **Displaying the validation of the selected image**

Positive result

In the case of validation with a score of at least 70%, a green icon and the "AR View" button will be displayed (Fig. 10.7).

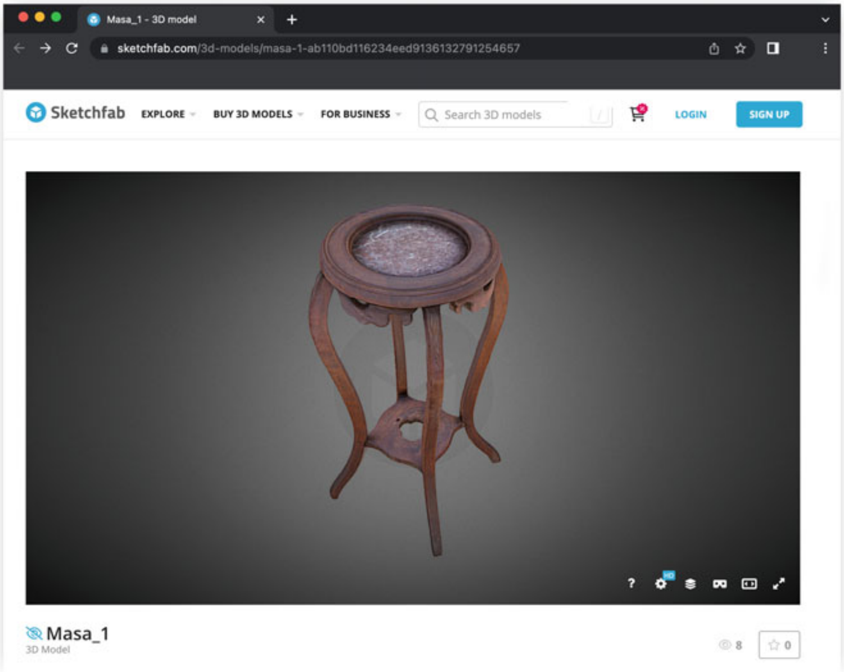


Fig. 10.2 The 3D model uploaded in Sketchfab (photo by M. Hodea)

Negative result

In the case of a score below of 70%, only a red icon will be displayed. The AR process will be disabled (Fig. 10.8).

(4) **The AR View for the validated image**

The object in context, augmented with historical explanations (Fig. 10.9)

The object in context, augmented with the representative 3D model (Fig. 10.10)

10.6.5 Implementation of the "FURNITURE" Application

For the implementation of the AI/ML process the AutoML framework was chosen (Google AutoML 2022), which is part of Vertex AI, the unified ML platform from Google (Google Vertex 2022). AutoML automates the ML model training to support developers in reducing time and effort related to ML cumbersome tasks, such as data preparation, choosing the right ML model, data parametrization and optimization.

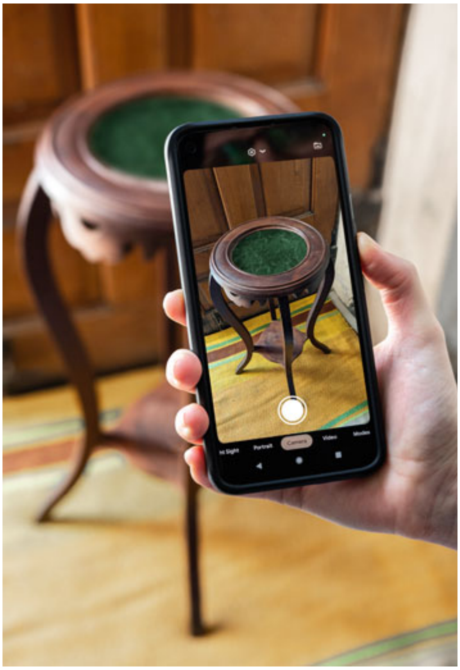


Fig. 10.3 Photographing a furniture object similar to the reference (photo by M. Moţăianu)

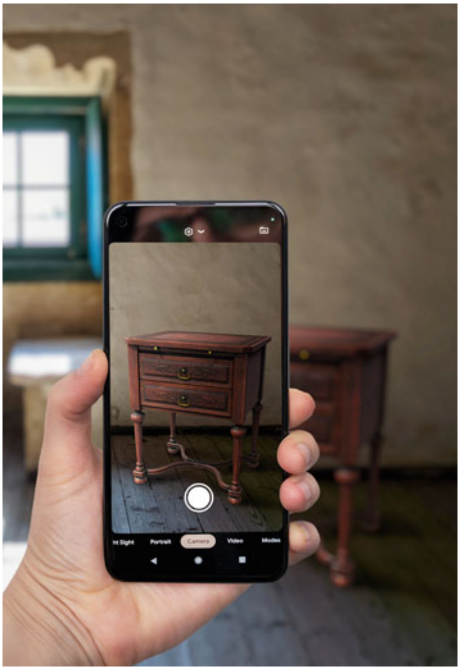


Fig. 10.4 Photographing a furniture object different from the reference (photo by M. Moţăianu)

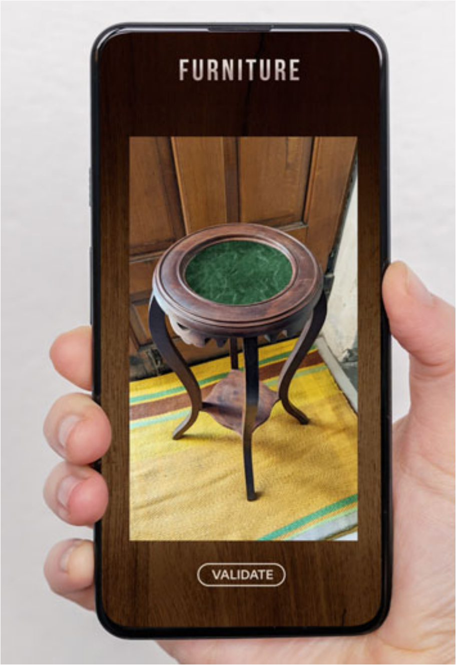


Fig. 10.5 Initiation of the ML process for a selected furniture image similar with the reference (photo by M. Moţăianu)

Fig. 10.6 Initiation of the ML process for a selected furniture image different from the reference (photo by M. Moţăianu)

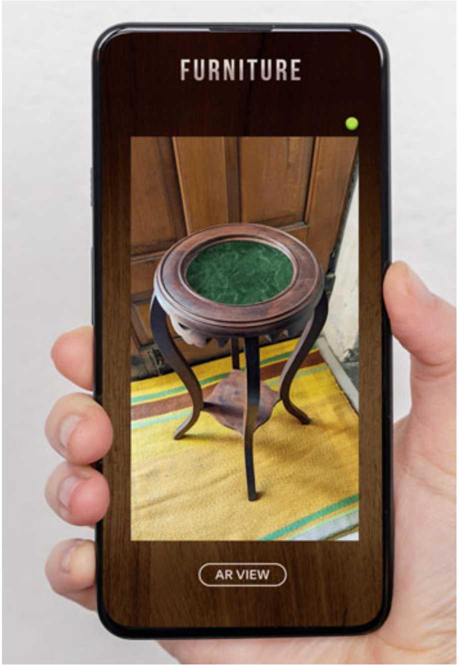


Fig. 10.7 Display of the information related to an image validated as similar with the reference (photo by M. Moţăianu)

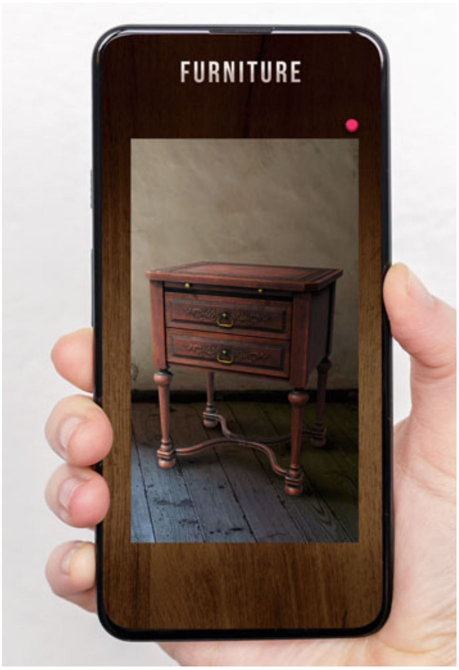


Fig. 10.8 Display of the information related to an image not validated as similar to the reference (photo by M. Moţăianu)

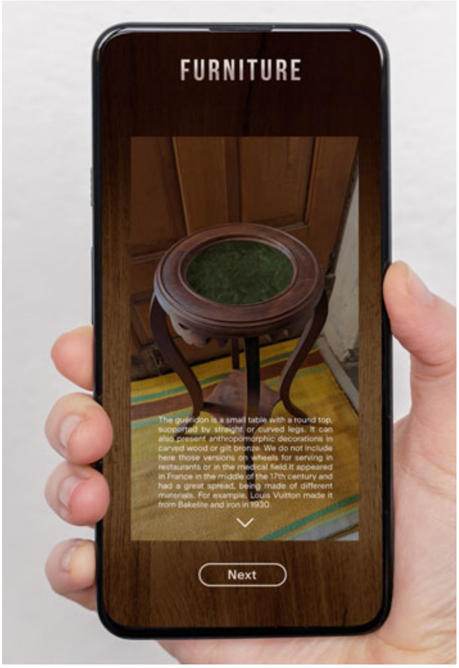


Fig. 10.9 AR view, the object in context, augmented with historical explanations (photo by M. Moţăianu)

Fig. 10.10 AR view, the object in context, augmented with the representative 3D model (photo by M. Moţăianu)

AutoML was used to perform a supervised learning training on the Vertex AI platform for an image classification model. The classification was related to a pre-defined furniture style, in our case the Guéridon table, and was performed on an image dataset of 180 images.

The most important step was the data preparation: (a) images generated from the 3D scans were used to create an annotated dataset (Google Vertex Annotations 2022), in our case, a label annotation was necessary to specify the Guéridon style; (b) the images and their labels were imported in a CSV file which was uploaded into the Vertex AI platform.

The dataset was automatically split by Vertex AI platforms as follows: 80% of images were used for training; 10% for validation and 10% for testing. After an evaluation of the model accuracy, the custom ML model was deployed as a REST API, to be used by the "FURNITURE" mobile application, in order to decide if a furniture image falls into the pre-defined category.

For the implementation of the AR and the user experience processes, the Wikitude AR framework was chosen. The authors considered from their previous research (Gheorghiu and Ştefan 2019; Gheorghiu et al. 2020, 2021) that the Wikitude AR framework is a fully-featured AR platform. Although it is a commercial platform, i.e. with incurring costs, it has the advantage of offering support for all AR paradigms. Wikitude Studio Editor is a web AR and content management platform for developing, managing, and publishing AR projects (Wikitude Studio 2022). The Studio Editor was used to upload the AR image target (WTC file) corresponding to the Guéridon style. The WTC file was used to attach the augmentations programmatically with Wikitude SDK and the JavaScript API (Wikitude SDK 2022).

10.7 Conclusions

This chapter has investigated a fusion of AI/ML with AR by designing and implementing an innovative mobile AR application named "FURNITURE". The application was designed for a new AR experience, different from the one previously experimented by the authors, which consisted in triggering the augmentations for objects or images identical to the reference images (the "targets"). The purpose of the current research was to leverage ML to detect images similar to the target ones and thus to trigger the AR process only for the images falling into a certain category, in our case, the Guéridon style.

Thus, the application can be used for an unlimited number of images, as the AR process is conditioned by the similarity with a reference image, defined by a minimal score of 70%. Furthermore, the application is open for future extensions, by expanding the object list that can be recognized and classified.

The ML model was trained using an automated supervised learning process with Google AutoML framework (Google AutoML 2022), and the classification results were tested as satisfactory for the purpose of the current research. A custom training can be addressed in the future.

Through the proposed application, the images sent by users to be recognized by the ML model can be collected and annotated to improve the ML model, provided that the users give their consent to participate in this continuous improvement process.

An application that allows a fusion between two powerful technologies, AI and AR, can enhance the pedagogical outcome, allowing the identification of any number of furniture objects augmented with a series of information with cultural and commercial value. This application can be widely used, from design or art history studies to the identification of objects on the antique or modern objects markets.

The application has the ability to develop a user's area of expertise through its action of adding new case studies, in case it is performed by students, teachers, or simple users.

In conclusion, we consider that the "FURNITURE" application provides a practical use (i.e., not only for demonstrative purposes), and therefore, has the potential for public appeal.

Acknowledgements The authors thank Professor Vladimir Geroimenko for the kind invitation to contribute to this book. Thank also to the art students who participated in the testing of the application. Last, but not least, many thanks to M. Bogdan Căpruciu for the useful comments and to Mrs. Cornelia Cătuna for the editing of the text.

References

3dmdb (2022) https://3dmdb.com/en/3d-models/Antique-Furniture/. Accessed Oct 2022

Adobe Photoshop (2022) https://www.adobe.com/ro/products/photoshop/. Accessed Oct 2022

Agisoft (2022) https://www.agisoft.com/. Accessed Oct 2022

Android UWB (2022) https://source.android.com/docs/core/connect/uwb. Accessed Oct 2022

Apple (2022) https://developer.apple.com/documentation/arkit. Accessed Oct 2022

Apple ARKit SDK (2022) https://developer.apple.com/augmented-reality/. Accessed Oct 2022

Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001) Recent advances in augmented reality. IEEE Comput Graphics Appl 21(6):34–47

Barceló JA (2006) Automatic archaeology. Bridging the gap between virtual reality, artificial intelligence and archaeology. In: Cameron F, Kenderdine S (eds) Theorizing digital cultural heritage: a critical discourse (Cambridge, MA, 2007; online edn, MIT Press Scholarship Online, 22 Aug 2013). https://doi.org/10.7551/mitpress/9780262033534.003.0023

Barceló JA (2009) Computational intelligence in archaeology. State of the art. In: Frischer B, Webb Crawford J, Koller D (eds) Making history interactive. Computer applications and quantitative methods in archaeology (CAA). Proceedings of the 37th international conference, Williamsburg, Virginia, USA, 22–26 March (BAR International Series S2079). Archaeopress, Oxford, pp 11–21

Barceló JA (2014) Visual analysis in archaeology. An artificial intelligence approach. In: Elwa A (ed) Morphometrics for nonmorphometricians. Lecture notes in earth sciences, vol 124. Springer, Berlin, pp 93–156, 51–101. https://doi.org/10.1007/978-3-540-95853-6_5

Bassett (2022) https://www.bassettfurniture.com/blog/furniture-style-guide.aspx. Accessed Oct 2022

Bishop CM (2006) Pattern recognition and machine learning (information science and statistics). Springer, New York

Bluetooth BLE (2022) https://source.android.com/docs/core/connect/bluetooth/ble. Accessed Oct 2022

Bogdanovych A, Rodriguez-Aguilar JA, Simoff S, Cohen A (2016) Authentic interactive reenactment of cultural heritage with 3d virtual worlds and artificial intelligence. Appl Artif Intell 24(6):617–647

Brosa (2022) https://www.brosa.com.au/blog/furniture-design-trends/. Accessed Oct 2022

Cgtrader (2022) https://www.cgtrader.com/free-3d-models/furniture. Accessed Oct 2022

Cipresso P, Giglioli IAC, Raya MA, Riva G (2018) The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Front Psychol 9:2086. https://doi.org/10.3389/fpsyg.2018.02086

Collins C, Dennehy D, Conboy K, Mikalef P (2021) Artificial intelligence in information systems research: a systematic literature review and research agenda. Int J Inf Manag 60. https://doi.org/10.1016/j.ijinfomgt.2021.102383. https://www.sciencedirect.com/science/article/pii/S0268401221000761

DesignConnected (2022) https://www.designconnected.com/. Accessed Oct 2022

Encyclopedia Britannica/AI (2022) https://www.britannica.com/technology/artificial-intelligence. Accessed Oct 2022

Encyclopedia Britanica/Guéridon (2022) https://www.britannica.com/technology/gueridon. Accessed Oct 2022

Foyr (2022) https://foyr.com/learn/best-interior-design-magazines/. Accessed Oct 2022

Furniture, Styylish History (2022) A guide to antique furniture leg styles. https://styylish.com/furniture-leg-styles/. Accessed Oct 2022

Gartner Report (2022) https://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017. Accessed Oct 2022

Gheorghiu D, Ştefan L (2019) Immersing into the past: an augmented reality method to link tangible and intangible heritage. Past for the future and future for the past: preservation and promotion of the world heritage sites, Sighisoara, Romania

Gheorghiu D, Ştefan L, Moţăianu M (2020) Augmented reality in environmental humanities education. In: Geroimenko V (ed) Augmented reality in education, a new technology for teaching and learning. Springer series on cultural computing. Springer, Cham, pp 353–366. https://doi.org/10.1007/978-3-030-42156-4

Gheorghiu D, Ştefan L, Hodea M (2021) Gestures and re-enactments in a hybrid museum of archaeology: animating ancient life. In: Geroimenko V (ed) Augmented reality in tourism, museums and heritage, a new technology to inform and entertain. Springer series on cultural computing. Springer, Cham, pp 153–172. https://doi.org/10.1007/978-3-030-70198-7

Grabcad (2022) https://grabcad.com/library/tag/furniture. Accessed Oct 2022

Google ARCore SDK (2022) https://developers.google.com/ar. Accessed Oct 2022

Google ARCore Fundamentals (2022) https://developers.google.com/ar/develop/fundamentals. Accessed Oct 2022

Google Assistant (2022) https://assistant.google.com/. Accessed Oct 2022

Google AutoML (2022) https://cloud.google.com/automl. Accessed Oct 2022

Google AutoML Deployment (2022) https://cloud.google.com/vision/automl/docs/deploy. Accessed Oct 2022

Google ML (2022) https://developers.google.com/ar/develop/java/machine-learning. Accessed Oct 2022

Google MLKit (2022) https://developers.google.com/ml-kit. Accessed Oct 2022

Google Vertex (2022) https://cloud.google.com/vertex-ai. Accessed Oct 2022

Google Vertex Annotations (2022) https://cloud.google.com/vertex-ai/docs/datasets/create-annotation-set. Accessed Oct 2022

Google Vertex Overview (2022) https://cloud.google.com/vertex-ai/docs/training-overview. Accessed Oct 2022

Gualandi ML, Gattiglia G, Anichini F (2021) An open system for collection and automatic recognition of pottery through neural network algorithms. Heritage 4(1):140–159. https://doi.org/10.3390/heritage4010008

IBM CV (2022) https://www.ibm.com/topics/computer-vision. Accessed Oct 2022

Instant Meshes (2022) https://www.blendernation.com/2015/11/16/instant-meshes-a-free-qaud-based-autoretopology-program/. Accessed Oct 2022

Janiesch C, Zschech P, Heinrich K (2021) Machine learning and deep learning. Electron Markets 31:685–695. https://doi.org/10.1007/s12525-021-00475-2

Konečný R, Syllaiou S, Liarokapis F (2016) Procedural modeling in archaeology: approximating ionic style columns for games. In: 2016 8th international conference on games and virtual worlds for serious applications (VS-GAMES), pp 1–8. https://doi.org/10.1109/VS-GAMES.2016.7590358

Lee JRH, Wong A (2020) AEGIS: a real-time multimodal augmented reality computer vision based system to assist facial expression recognition for individuals with autism spectrum disorder. Vision and Image Processing Lab. https://www.resna.org/sites/default/files/conference/2020/NewEmergingTechnology/73Lee.html

Meshmixer (2022) https://www.meshmixer.com/. Accessed Oct 2022

Niemi H (2021) AI in learning: preparing grounds for future learning. J Pac Rim Psychol 15. https://doi.org/10.1177/18344909211038105

Open3dmodel (2022) https://open3dmodel.com/3d-models/antique-furniture. Accessed Oct 2022

Pokemongo (2022) https://pokemongolive.com/en/. Accessed Oct 2022

Pokemongo Succes (2022) https://gecon.es/pokemon-go-success. Accessed Oct 2022

Popenici SAD, Kerr S (2017) Exploring the impact of artificial intelligence on teaching and learning in higher education. RPTEL 12:22. https://doi.org/10.1186/s41039-017-0062-8

Rauschnabel PA, Felix R, Hinsch C, Shahab H (2022) Alt F (2022) What is XR? Towards a framework for augmented and virtual reality. Comput Hum Behav 133:107289. https://doi.org/10.1016/j.chb.2022.107289

Samuel L (1959) Some studies in machine learning using the game of checkers. IBM J Res Dev 3(3):210–229

Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. SN Comput. Sci. 2:160. https://doi.org/10.1007/s42979-021-00592-x

Sketchfab (2022) https://sketchfab.com/. Accessed Oct 2022

Sketchfab Antique (2022) https://sketchfab.com/tags/antique-furniture. Accessed Oct 2022

Unity (2022) (https://unity.com/unity/features/arfoundation. Accessed Oct 2022

Vervoordt (2022) https://www.mapswonders.com/axel-vervoordt/. Accessed Oct 2022

Vuforia (2022) https://developer.vuforia.com/. Accessed Oct 2022

Wikitude (2022) https://www.wikitude.com/developer-overview/. Accessed Oct 2022

Wikitude ARCore (2022) https://www.wikitude.com/external/doc/expertedition/#how-does-wikitude-sdk-relate-to. Accessed Oct 2022

Wikitude Concepts (2022) https://www.wikitude.com/external/doc/expertedition/Concepts.html. Accessed Oct 2022

Wikitude Trackers (2022) https://www.wikitude.com/multiple-trackers-augmented-reality/. Accessed Oct 2022

Wikitude Studio (2022) https://www.wikitude.com/products/studio/. Accessed Oct 2022

Wikitude SDK (2022) https://www.wikitude.com/products/wikitude-sdk/. Accessed Oct 2022

Part III
Augmented Reality and Artificial Intelligence in Medicine, Healthcare, and Physical Activity

Chapter 11
Meta-patients: Using Mixed Reality Patients and an AI Framework for Simulating Life-Like Clinical Examinations

Gary Grant, Rob Burton, Eileen Grafton, Daniel Della-Bosca, Robert Ditcham, and Louise Humphreys

Abstract This chapter addresses an augmented learning experience created for the Griffith University School of Nursing and Midwifery in 2022. An interdisciplinary research team from Griffith University (Australia) deployed the first iteration of an application for students in response to the difficulties imposed through the previous two years. The impacts of the COVID pandemic bought challenges to the Bachelor of Nursing program, particularly in relation to student competency in the physical assessment of patients, through objective structured clinical examinations. This pilot study introduced life-like, simulated patients, designed and rendered within Unreal Engine to the students. The patients were accessible through cross platform applications, including mixed reality devices. Students were also able to interact with patient information communicated using the AI framework afforded by Microsoft PowerApps all packaged in a bespoke SharePoint site. Student participants were interviewed as part of the development process and approved of Augmented and

G. Grant
School of Pharmacy and Medicine, Griffith University, Gold Coast, Australia
e-mail: G.Grant@griffith.edu.au

R. Burton
Department of Nursing, Midwifery and Health, Northumbria University, Newcastle, UK
e-mail: rob.burton@northumbria.ac.uk

E. Grafton
School of Nursing and Midwifery, Griffith University, Logan, Australia
e-mail: e.grafton@griffith.edu.au

D. Della-Bosca (✉) · R. Ditcham
Queensland College of Art, Griffith University, Gold Coast, Australia
e-mail: d.della-bosca@griffith.edu.au

R. Ditcham
e-mail: r.ditcham@griffith.edu.au

L. Humphreys
Technical Partners Health, Griffith University, Gold Coast, Australia
e-mail: louise.humphreys@griffith.edu.au

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_11

Mixed Reality as successful platforms for the deployment of the simulated patient scenarios within Microsoft Teams and available through a mobile application and mixed reality device.

11.1 Introduction

For over 20 years, Griffith University (Australia) has provided an eight-course Bachelor of Nursing program in Singapore for local Registered Nurses (RNs) to transition from a Diploma of Nursing to bachelor level degree. Biosecurity measures introduced at the start of the COVID-19 pandemic in February 2020, meant all teaching was moved online for the foreseeable future. Responding to the necessity to continue teaching, two of the authors (Professor Robert Burton and Ms Eileen Grafton) sought methods to pivot to online teaching and assessment.

Teaching methods were adapted quickly using pre-existing learning management systems (LMS) and virtual lectures and clinical skills laboratories replaced in person teaching (Grafton et al. 2021). Assessment posed a more complex problem as OSCE's (objective structured clinical examinations), were normally conducted with a 'real' simulated patient, but biosecurity measures introduced in Singapore in February 2020 made it impossible for students to access even volunteer patients. A 'virtual' patient in the form of images taken from various skills videos was used. Students, using Collaborate within the LMS, were able to augment the images in real time using the interactive whiteboard.

OSCEs for the Clinical Health Assessment course were able to be completed online in 2020, through the innovative use of available technology although the reality of the experience was somewhat rudimentary (Grafton et al. 2021).

To improve on the participant's sense of immersion, for the 2022 iteration of the course, a research project was initiated in 2021 to explore opportunities for crafting a mixed reality experience, combining simulated patients and a virtual learning environment. As an outcome of this project, an integrated application for OSCEs was developed, implemented and evaluated. The integrated application (MR App), was developed using a design-based research approach, informed by iterative feedback from students and staff during the development. The app was built using the Microsoft Power Platform. This application allowed for dynamic content to be displayed and designed for student control during use. Realistic virtual patients were constructed using *Metahuman Creator* (Metahuman Creator 2022) and *Unreal Engine* (Unreal Engine 5 2022). Their rendered outcomes were incorporated in the application with pen input to allow for student interaction indicting examination landmarks. The application contained all aspects of the clinical health assessment, including patient information, patient interaction, physical assessments, and clinical handover. Although the project sought to investigate an improvement to clinical examination in virtual terms, the investigative team explored and tested new opportunities for digital patient simulation in virtual, augmented, and mixed reality. The projects outcome is an integrated application, but the project explored opportunities for new areas of investigation for the development of digital simulated patients.

11.2 Augmenting the Design Process

As the projects goals were driven by pedagogy, the research methodology adopted for the project was design based (Wang and Hannafin 2005; The Design Based Research Collective 2003). Of note however, were the varying approaches and interpretations to a design-based methodology based on the individual experiences and discipline areas of the team members. A design-based methodology can incorporate multiple iterations of content design but retains a student centric enquiry. It seeks to investigate related problems, pose potential answers and not simply to interrogate one problem and find a solution.

Herbert Simon, in *The Sciences of the Artificial*, writes of problem representation. Simon writes that “every problem-solving effort begins with creating a representation for the problem—a problem space in which the search for a solution can take place” (Simon 1996, p. 108). This idea of problem representation is important and largely misinterpreted in design contexts for Simon continues to explain that the problem space is personally informed by the individual that addresses the problem. Each representation of a problem is unique to each of us. Herbert Simon is sometimes misquoted as introducing us to the idea of problem finding, which unfortunately oversimplifies the epistemological issue of addressing a problem based on prior understanding and a personal and professional frame of reference. Donald Norman too, suffers the same misquote. Norman frequently writes not of problem finding but finding the right problem (Norman 2013). A design-based methodology can be better utilised when it is scrutinised epistemologically. There is no one way of knowing.

The research project described here sought a novel solution to a problem that arose because of the impact of isolation on teaching and learning. The first iteration of online learning implemented by Robert Burton and Eileen Grafton in 2020, described a problem space defined by personal and professional experience. The second iteration, the research project begun in 2021, added a few more epistemological frames to the creative pool. Each frame of each observer helped to redefine the problem space. Simon writes, “what constitutes novelty depends on what knowledge is already in the mind of the problem solver and what help is received from the environment in adding to this knowledge” (Simon 1996, p. 105). The problem space originally conceived, expanded quickly over the projects beginnings as a result of the creative collaboration between authors. Simon argues that design allows diverse disciplinarians to communicate using a shared language, to perceive the common creative activity with which they are engaged (Simon 1996). The problem goal was common, this concerned the need to effectively simulate a patient for a nursing student. The manner in which each member of the project team understood the goal could be argued to be fairly similar. The manner in which each member of the project team perceived the problem space was understandably quite different. And this is what gives rise to creative activity. Each member of the project team negotiated the problem space differently (represented the problem differently for themselves) but communicated their perspective to other team members as the experimentation unfolded and as the project evolved. The problem space from a nursing perspective sought to mirror the strict requirements of clinical examination and any new virtual solution needed to be performed and assessed in real time, as directed by the Singapore Nursing Board and the Nursing executive in Australia (Grafton et al. 2021). Those from the School of Design saw their problem space as one requiring visual fidelity and a negotiation of presence. That is, presence in the sense of perceptual realism and in the sense of presence as transportation (Lombard and Ditton 1997), the notion that the viewer is transported to a 'believable' spatial construct.

Relevant technologies that were experimented with, tested, and incorporated into the problem space in turn informed the problem space. Norman writes, "new technologies change the meanings of things. And creative people continually change how we interact with our technologies and one another" (Norman 2013, p. 130). As educators, we may strive for a pedagogically driven research approach. However, we also must negotiate the influences of the technologies we experiment with so that we can permit the affordances granted by the technologies to help shape the problem space in turn. The Design Research methodology when investigating new technologies affords serendipity. The problem space can expand and allow for greater creative interaction between collaborators if there is room for serendipity. Pedagogy and serendipity may seem to be terms that are not complementary to each other, but it is a valuable outcome of the collaborative experience exploring technology in general and in particular educational technologies. Regarding this project, the educational technologies included Microsoft SharePoint, Microsoft PowerApps, Microsoft Power Virtual Agents for Teams, and Microsoft Teams. The project also utilized technologies from outside the educational technology paradigm, including Character Creator, Autodesk 3DS Max, Unreal Engine and Metahuman Creator. The technologies that were co-opted for use in building new scenarios for teaching and learning required considerable experimentation as some were completely new to the project team. Thomas Reeves writes of the technological innovations as included in the development of the solution step of the Design Research methodology, see Fig. 11.1. The affordances of these technological innovations mentioned above necessarily influenced the research project but so did the technological affordances of the mediums of deployment.

Experimentation with technologies to deploy the simulated scenarios, (personal computer, mobile phone and mixed reality headset), at the development and testing stages created unique opportunities that in turn helped reshape the project as a whole. For the course delivered in 2020, in the deployment of virtual patient in image form, augmentation of the scenario was simply undertaken by the student drawing on the screen and overlaying the virtual patient with task information within the blackboard collaborate virtual room (Grafton et al. 2021). The opportunity for the simple success of low fidelity augmentation utilizing web-based software, was afforded by

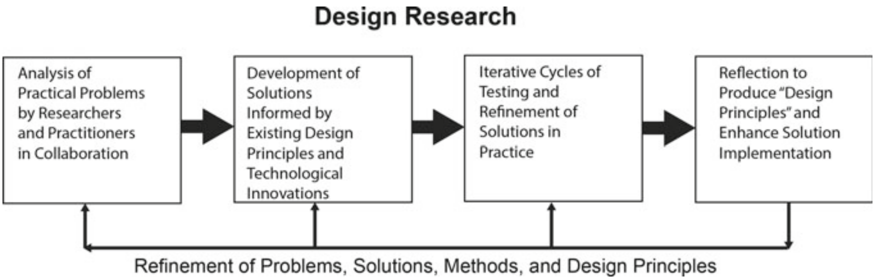


Fig. 11.1 Thomas C Reeves, design research approaches in educational technology research (Reeves 2006). Copyright (2006) From (Design research approaches in educational technology) by (Thomas C Reeves). Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc

a component within the learning management system. This affordance, a forced but serendipitous constraint, re-defined the problem space in the next iteration to include other web-based software platforms for deployment.

The current project would not have benefited from the rigor of some design-based research methods, even hypothetical explorations, were it not for collaborative play with technologies, The implementation of new technologies, such as mixed reality technologies for instance, continue to be referenced in design-based research (Cowling and Birt 2018), but Reeves's model is cited in an abbreviated form, which is problematic as the dynamic nature of creative collaboration is not necessarily considered. Reeves noted this himself and credited Ann Brown and Alan Collins for the characteristics of *design experiments* as addressing complex problems in real contexts in collaboration with practitioners (Brown 1992; Collins 1992). Reeves although promoting a strict set of procedures for design-based research, left space for collaboration and for development of solutions informed by technological innovations (Reeves 2006). Unfortunately, now this model can sometimes be abbreviated to the following: analysis of problem—development of solution—testing and refinement—reflection on implementation (Cowling and Birt 2018). Technological Innovation in the context of this project, in simple terms could be explained as follows: experiment with the technologies and see what they were not intended to do.

11.3 The Problem of the Simulated Patient

At the onset of the project in 2021, the concepts used to discuss the requirements of the project focused on the idea of virtual. The *virtual patient* and *virtual rooms* were discussed by Professor Burton and Ms Grafton (Grafton et al. 2021) in their reflection on the initial iteration using Blackboard Collaborate. The project name was determined to be AV/VR, as a response to the teaching practices using a virtual patient. It should be noted that there are domain specific ideas of virtual. Clinical

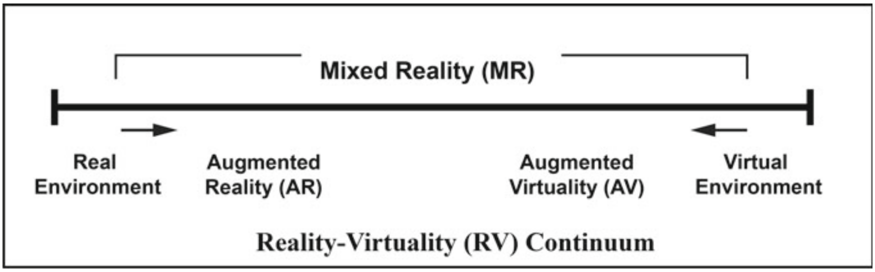


Fig. 11.2 Virtuality continuum (Milgram and Kishino 1994)

health has long used simulated patients which are *real* persons. A simulated patient that is digitally created and communicated is simply a virtual patient regardless of the mode of deployment. In the domain of digital design, every model, scene, and image created in software is virtual as its very existence is but information. Sometimes we consider things more real than virtual and sometimes we consider things more virtual than real. It is largely a cognitive position (Chalmers 2022).

Paul Milgram and Fumio Kishino are credited as developing the often-cited *Virtuality Continuum* diagram illustrated in Fig. 11.2 (Milgram and Kishino 1994).

This diagram, originally used by Milgram and Kishono to explain the classes of mixed reality interfaces, has become more generally useful in exploring an entire spectrum of ideas relevant to mixed reality and not limited to digital technologies. Mixed reality, however, is only comparable as an ideology when compared to similar continuums such as the eXtended Reality continuum (XR). eXtended Reality can be understood as experiences that are made possible through the use of XR software coupled with specialized hardware including, but not limited to, head-mounted displays (HMDs), sensors, and motion controllers (Herur-Raman et al. 2021), see Fig. 11.3.

Investigations for the AV/VR project certainly involved the discussion, trial, and experimentation of head-mounted displays, technologies but they were never considered a requirement for the project's success. The idea of mixed reality however, and the accessibility of digitally created media via whatever technologies were most appropriate for the student experience were important. The Microsoft Hololens HMD was trailed partly for its fidelity in augmenting digital content, but especially because of its integration with Microsoft Sharepoint spaces which permits, web based content to be easily accessed. It should be noted that although the definitions of AR as a set of functional interactive technologies continue to be applied to particular devices and experiences (Herur-Raman et al. 2021). The term Mixed Reality for device and experience is used by the research team and others similar researchers (Wagner et al. 2009; Cowling and Birt 2018). Mark Billinghurst also recognises the popular shift in speaking about Mixed Reality and states that "almost any display that combines real and virtual imagery is a Mixed Reality experience" (Billinghurst 2017).

The initial stages of the AV/VR project involved discussions concerning what was the nursing team desired in the simulated scenarios and a dialogue to help the design

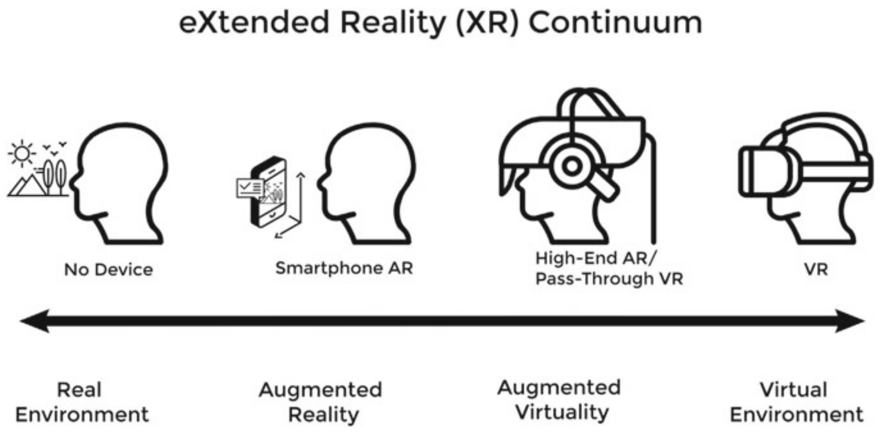


Fig. 11.3 Extended reality continuum (Herur-Raman et al. 2021) CC by 4.0 https://creativecommons.org/licenses/by/4.0/

team understand the clinical procedures involved. The entire team in turn adapted the scenarios of the nursing team to produce simulation storyboards for everyone to contribute to, see Fig. 11.4. The initial virtual patients were trialled in both Autodesk Character Generator and Character Creator by Reallusion. Character Creator was selected as the fidelity of the 3-dimensional models is high and appropriate for subsequent rigging.

The models produced by character creator were quite appropriate as the texture maps were of high quality and models were able to be rigged and exported.

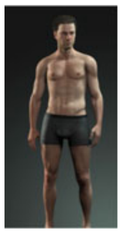
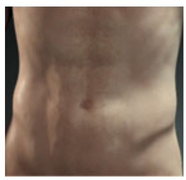
Peripheral Vascular assessment, the student needs to: • Note the overall colour, warmth of the skin • locate / identify each pulse site on arms and legs and palpate the pulses at these sites • Assess peripheral refill on fingers and toes • Summarise the findings	Interaction: Web based, mouse over spots. Student potentially draws on screen	
GIT / abdominal The student needs to visualise the abdomen, noting appearance contour etc. • Identify the 4 quadrants • Auscultation (with stethoscope), palpate and percuss over the 4 quadrants identifying the underlying organs in each area and discussing findings • Summarise the findings	Interaction: Web based, mouse over spots. Student potentially draws on screen	

Fig. 11.4 AV/VR project, excerpt from initial storyboard

Metahuman Creator was trialled shortly after and the ease of use in customising features proved to be of great benefit. The potential for the nursing team to be able to continue creating virtual patients and then pass the models on to the design team for subsequent rigging was favoured over other methods of character creation. The Metahuman characters were able to be designed to suit the written histories of their patient file inclusive of age and ethnic features. As the AV/VR project was conducted with the nursing program in Singapore, the simulated patient histories included peoples of Singaporean and Malaysian origin. *Mary Leong* illustrated in Fig. 11.5 demonstrates the design of facial characteristics, to match the existing simulated patient file. A total of three simulated patients were created for the project, some had several iterations of development using Character Creator and Metahuman Creator. The patients included *Mary Leong* (illustrated in Fig. 11.5), *Irfan Bin Rahman* and *Rasheeda Khalid*.

The names and detailed patient histories are important not just from the standpoint of clinical training but to instil presence. The high level of realism afforded by Metahuman Creator assists with Lombard and Ditton's assertion of presence as realism, "the degree to which a medium can produce seemingly accurate representations of objects, events, and people—representations that look, sound, and/or feel like the *real* thing" (Lombard and Ditton 1997). Realism here can be understood as perceptual realism, but also as social realism, using the personal characteristics of the simulated patient to build a detailed patient history through familiarity and association. The repeated use of name instead of referring to patient x, builds familiarity and believability. We are prepared to believe in the narrative.

Fig. 11.5 Mary Leong, simulated patient AV/VR project, 70 years old, 158 cm, 52 kg

Fig. 11.6 Mary Leong, simulated patient AV/VR project

The next iteration of the development of simulated patients involved export to Unreal Engine 5, character rigging and posing, and the creation of garments appropriate for the clinical experience, illustrated in Fig. 11.6. The garments are carefully tailored to match the Queensland Health hospital gown so that they would be appropriate for ongoing use and would be familiar to Nursing students of Griffith University.

The export to Unreal Engine was also important to be able to merge a scene file created in Autodesk 3DS Max of a complete hospital ward, tailored to Queensland Health standards. The scene and characters were merged so that each simulation could be contextualised for necessary steps in each clinical simulation. Referring again to the Lombard and Ditton's concept of presence, the hospital ward addresses their point concerning realism as transportation, specifically the 'you are there' construct in which the user is transported to another place (Lombard and Ditton 1997). Lombard and Ditton admit this concept is perhaps the oldest version of presence, having roots in oral and literary history. The idea of presence as transportation was also conceptually introduced in different but comparable terms by Samuel Taylor Coleridge as "willing suspension of disbelief" (Coleridge 1817). This concept after Coleridge was adopted by the medium of cinema in the early twentieth century and continues to be advocated in the newer mediums of virtual and augmented reality (Slater and Usoh 1993; Wagner et al. 2009).

The final stage of character development in Unreal engine involved the posing of the simulated patient for each of the clinical assessment tasks. Figure 11.7 shows the simulated patient Irfan Bin Rahman posed for abdominal examination.

The level of perceptual realism was retained to an acceptable degree within Unreal Engine, but only by preserving a high level of detail (LOD) inclusive of texture map and hair grooms which are effective for use within Metahuman Creator and Unreal Engine but not exportable to other applications (see Fig. 11.8).

Fig. 11.7 Irfan Bin Rahman, simulated patient AV/VR project

Fig. 11.8 Irfan Bin Rahman, profile, simulated patient AV/VR project

The outcomes from Unreal Engine met the project teams goals as the exported renders from Unreal Engine were to be then incorporated into the browser-based application, Microsoft Power App. It is possible to deploy entire scenarios for student access in Unreal Engine, however the lessons learned from the previous two years spent in imposed isolation, sought platform independent scenarios requiring no specialised hardware or specialised technicians to facilitate the scenarios. There is evidence that simulation for nursing is exploring a range of technological approaches (Vaughn et al. 2016; Hauze et al. 2019; Kyeng-Jin et al. 2021) and a recent meta-review of XR simulation states, "medical simulation and specifically eXtended Reality technologies including Virtual and Augmented Reality are being "adopted by many healthcare institutions and will be essential components of the post-pandemic future of medical education" (Herur-Raman et al. 2021, p. 12). We argue that the high-end technologies for XR are not always appropriate to the task. This is supported by other investigators as the focus should be on student/ user experience and not technology (Hauze et al. 2019; Cowling and Birt 2018). This project sought to purposefully adopt tools and technologies particularly browser-based technologies so that the scenarios could be easily experienced by students with easily available hardware.

11.4 Interacting with Virtual Patients

An important goal of this project was that the virtual clinical simulations should be easy for students to interact with. Browser-based simulations permitted access by students with available hardware, and the Microsoft ecosystem permitted integration with our learning management systems. To begin with, the project team explored and tested several applications within the Microsoft SharePoint framework, including SharePoint Spaces which is a browser based Mixed Reality platform that can integrate easily with the Microsoft HoloLens.

The project team desired that a complete simulated experience with virtual models deployed using Mixed Reality devices would be appropriate. The hardware independence of SharePoint Spaces seemed to support the idea that elements of the simulated scenarios could be accessed using a Microsoft HoloLens if available, or if not, on mobile devices that the students possessed. The pedagogical necessity was that all aspects of the simulated scenarios would be accessible on a hardware platform most appropriate and available to each student.

To be able to deploy a scenario where the student could interact with a virtual simulated patient overlaid onto a real and familiar environment, was what the project team sought. And in particular, the project team sought a high-fidelity augmented reality experience.

Our efforts using SharePoint Spaces and the Microsoft HoloLens, required three dimensional models of a particular file size with specific textures however, and so attaining high levels of realism proved out of reach.

The highly realistic simulated patients already created in Metahuman Creator proved to be exactly what was required, but the use of Metahumans is restricted to Unreal Engine. Deploying simulated scenarios through Unreal Engine using a high-end computer is certainly achievable but did not comply to the project goal of platform independence. The high-fidelity characters already created however afforded opportunity. The still images of the characters were shown to a focus group of Nursing students, who determined that the images alone were effective in building a rapport with a simulated patient. The next stage of the project involved posing the characters within their virtual environments and rendered compositions, two of these are illustrated in Figs. 11.6 and 11.7. The next step involved the integration of the still image compositions into the browser-based application, Microsoft Power Apps.

Microsoft Power Apps is a low-code application that allows for the creation of customised apps within a learning management system. The application permits the developer to link available data and develop a single, interactive tool which integrates seamlessly with Microsoft Teams for delivery.

The interactive elements available in Power Apps with relevance to this project included timer, text input, pen input, camera, measuring camera, video recording, audio recording, view in mixed reality, and markup in mixed reality. In addition, the ability to link to other Power Platform applications, permitted the development of a single contained tool with diverse capabilities that could be delivered easily within Microsoft Teams. Microsoft Power Apps allowed for the creation of an interactive tool to deliver the sequence of events required by each student to complete. The application allowed for identification of the user, user permission control, text content for the relevant scenario, individual timing of student activity, and transitions which allowed for self-paced progression through the tool. The interactive pen function allowed for students to demonstrate important concepts on relevant selected images. The app allowed for content and images to be pulled from associated databases thereby simplifying scenario build and editing times. Embedding the Power Apps within Microsoft Teams allowed for simultaneous use of the chat function, private breakout rooms, video meetings, and session recordings for playback and review.

Figures 11.9, 11.10, 11.11 and 11.12 show some screen interfaces from the completed Microsoft PowerApp.

The interactive pen tool illustrated in Fig. 11.11, permitted the students to interact with the simulated virtual patient and perform relevant surveys as required. This was a feature of clinical examinations developed by Professor Burton and Ms Grafton in 2020 out of necessity (Grafton et al. 2021). This serendipitous outcome was refined

Fig. 11.9 Clinical health assessment application start screen

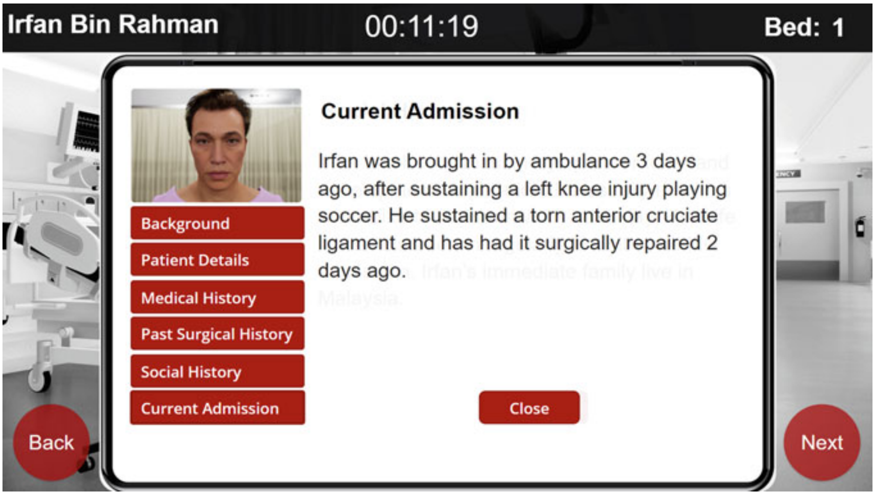


Fig. 11.10 Clinical health assessment application, patient details Irfan Bin Rahman

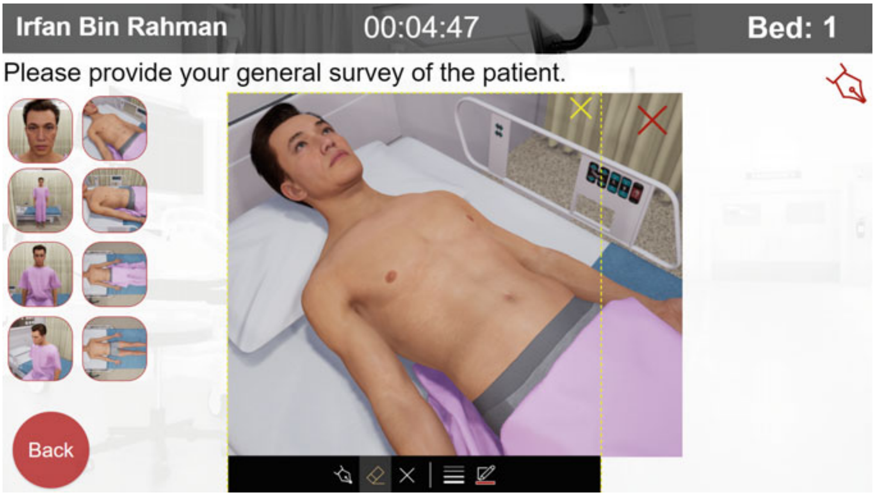


Fig. 11.11 Clinical health assessment application, patient survey

for use in the Microsoft PowerApp and the interactive tools proved successful with student participants.

David Chalmers in *Reality+*, poses an alteration to Milgram and Kishino's *reality-virtuality continuum.* Chalmers writes that "the continuum is misnamed, because it

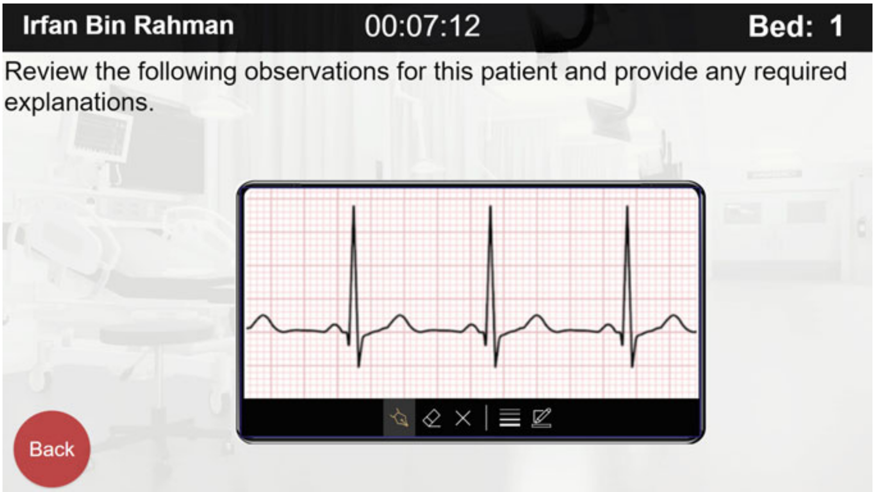


Fig. 11.12 Clinical health assessment application, vital signs

bakes in the premise that virtuality is opposed to reality" he offers the *physicality-virtuality continuum* (Chalmers 2022). The virtual patient in Fig. 11.11, with superimposed actual drawings by students during examination, adheres to Chalmers's definition of *augmented virtuality*, where a virtual world serves as the basis, augmented by physical objects (Chalmers 2022). The drawings and notations by students in this case, the physical objects, tangible artefacts resulting from interaction with the virtual patient.

Lombard and Ditton's "final conceptualization of presence involves social responses of media users not to entities (people or computer characters) within a medium, but to cues provided by the medium itself" (Lombard and Ditton 1997). This they refer to as *Presence as medium as social actor* (Lombard and Ditton 1997). Dr Grant, during the course of the project, created AI powered conversational bots to provide another layer of interaction within the clinical health assessment application. These bots like any trained chatbot, interact in real time and fill traditionally social roles (patient, teachers). Lombard and Ditton write, "basic social cues exhibited by the medium lead users to treat the medium as a social entity" (Lombard and Ditton 1997).

Within the Microsoft ecosystem the AI conversational bots are called Power Virtual Agents, these are conversational bots powered by natural language processing. These AI-powered bots are created to play scenario patients and other healthcare professionals (as illustrated in Fig. 11.13), thereby allowing students to cross check provided information and gather more detailed information during the activity. Trigger phrases are entered for a range of important topics and the associated response messages pre-populated. The bot allows for non-linear transition through the scenario and creates a much more authentic demonstration of the students' skills and capabilities. The deployment of a bot within Microsoft Teams also simplifies

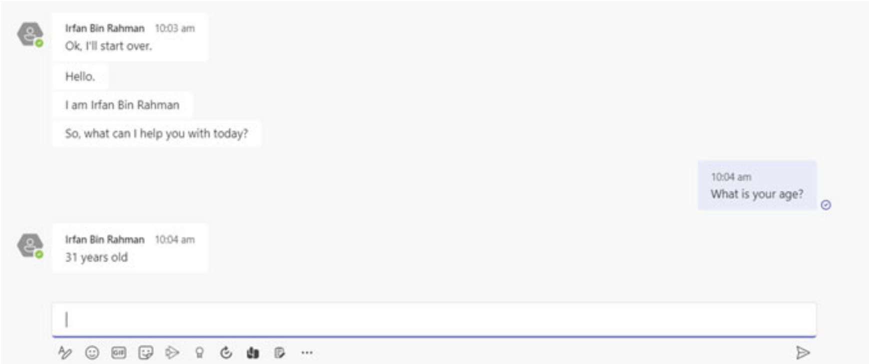


Fig. 11.13 Clinical health assessment application, chatbot excerpt

integration and permission control. This functionality was trialled as part of the preliminary project and will form an important component of successive versions of the application.

There is of course dissent in the VR/MR/XR research community as to exactly what comprises presence or whether perceptual presence is an appropriate indicator of believability. Stefan Weber, David Weibel and Fred Mast, for example, counter perceived realism with the notion of *being there*. They "define *being there* as the allocation of attentional resources to the mediated world and the sensation of perceptually being surrounded by the Virtual Environment" (Weber et al. 2021, p. 2). This definition does seem true for the virtual environment that surrounds the viewer, for the experience skewed toward the virtual end of the *physicality-virtuality continuum of David Chalmers*. Chalmers does however support the psychological notion of immersion. He writes that there are many degrees of immersion and uses the example of a video game on computer screen occupying all of our attention in a *flow state* as psychological immersion (Chalmers 2022).

The AI conversational bot discussed above is but one component of an integrated set of technologies within the Microsoft PowerApp that negotiate a participant's investment in the experience. Regarding Lombard and Ditton's ideas of the illusion of non-mediation, it is quite relevant, as an AI conversational bot engenders familiarity. The experience with a conversational bot requires time and each question or answer that is understood as meaningful to the participant adds slowly to psychological immersion. Lombard and Ditton's concept of the illusion of nonmediation can be effectively illustrated with the conversational bot, "the medium appears to be transformed into something other than a medium, a social entity" (Lombard and Ditton 1997). In subsequent iterations of this project, it is expected that training the conversational bots and implementing their use in many varied scenarios will begin to instil for the students a familiarity. In extended simulations involving interaction with the bot/ simulated patient it is expected that with each instance of communication the acceptance of social entity will grow. Lombard and Ditton were careful to acknowledge this, stating that presence in this view, does not occur in degrees nor does it occur at once, it is the subjective feeling of a greater number of instants during the experience in which the illusion of nonmeditation occurs (Lombard and Ditton 1997).

11.5 Conclusions and Future Plans

A focus group, comprised of students that had undertaken the course in the prior year, was run to demonstrate the tool and collect feedback for refinement of the application. Not having been exposed to the application previously, this feedback on user experience provided instructive insights for refinements. Renderings of the simulated patients were shown to the focus group at this time, and it was found that the students were accepting of the images embedded within the test application. At the culmination of this project in 2022 and its deployment, students were surveyed in order to gain qualitative feedback on the clinical health application and their interactions with it. The team did not seek to measure presence, but the students' comments were positive. 92% of respondents reported that they felt involved or very involved with the visual aspects of the application. Some of the comments received included a student that wrote "I felt like I was talking to real patients after several practices" and "I'm impressed that I could actually see the abdominal, chest area, face and etc. clearly. It really helps me to visualise the actual sites properly for physical assessment". Not all comments were positive of course. The feedback did include criticisms of lack of clinical detail, i.e., the lack of bruises for instance. For subsequent iterations of the application and of the simulated patients, customisation of skin texture where necessary will be implemented.

The team has gained valuable IT experience during the project that is transferable and valuable. The team is now in the process of sharing outcomes through relevant research centres with Griffith University and through external publications and presentations by demonstrating the capabilities of the developed application and associated technologies. During the project, the team explored the benefits and limitations of Power Apps in Teams, Dataverse in Teams, Power Apps as part of the Power Platform, and other databases for data storage, addition, and editing. The application was built in both Power Apps in Team and Power Apps as part of the Power Platform. Although there are potential benefits of developing using Power Apps in Teams coupled with Dataverse in Teams, there is a reduction in available tools for development. These tools include inclusion of virtual objects, and augmented reality tools. Power Apps will become an important application for Hololens 2 soon. It will give Griffith Health and Design researchers the ability to develop custom AR applications at low cost that is customisable to the specific educational need.

Work with Microsoft during the project has made significant progress with regards to the use of AI-bots to augment and provide greater authenticity to simulations. The ability to integrate AI-bots directly into the developed PowerApp is now possible. In addition, the use of Azure cognitive tools provides the ability for speech to text and then text to speech. This removes the need for typing and opens the door for 52 language choices. The addition of this technology will create dynamic interactions between the student and the simulated patient. The aim of the project was to cross pollinate ideas and IT skills. The academics and technical staff in the team have shared important skills, which will likely transfer to others across discipline areas, thereby making the use of these technologies more sustainable.

Acknowledgements The project team acknowledges the support of the School of Health, Griffith, University for research funding which has helped realise this project.

References

Billinghurst M (2017) What is mixed reality? https://medium.com/@marknb00/what-is-mixed-reality-60e5cc284330. Accessed 20 Oct 2020

Brown AL (1992) Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. J Learn Sci 2:141–178

Chalmers DJ (2022) Reality+. Allen Lane, Dublin

Coleridge ST (1817) Biographia Literaria, Chapter XIV, Available via https://www.gutenberg.org/files/6081/6081-h/6081-h.htm#link2HCH0014. Accessed 17 July 2022

Collins A (1992) Towards a design science of education. In: Scanlon E, Shea TO (eds) New directions in educational technology. Springer, Berlin, pp 15–22

Cowling M, Birt J (2018) Pedagogy before technology: a design-based research approach to enhancing skills development in paramedic science using mixed reality. Information 9(29):1–15. https://doi.org/10.3390/info9020029

Grafton E, Elder E, Burton R (2021) Innovative strategies to maintain nursing students' academic continuity during the COVID 19 pandemic. J Appl Learn Teach 4(1):21–28. https://doi.org/10.37074/jalt.2021.4.1.7

Herur-Raman A, Almeida W et al (2021) Next-generation simulation—integrating extended reality technology into medical education. Front Virtual Reality 2. https://doi.org/10.3389/frvir.2021.693399

Hauze SW, Hoyt H et al (2019) Enhancing nursing education through affordable and realistic holographic mixed reality: the virtual standardized patient for clinical simulation. In Rea PM (ed) Biomedical visualisation, vol 1. Springer International Publishing, pp 1–13. https://doi.org/10.1007/978-3-030-06070-1_1

Kyeng-Jin K, Moon-Ji C, Kyu-Jin K (2021) Effects of nursing simulation using mixed reality: a scoping review. Healthcare 9(8), Article No. 947 https://doi.org/10.3390/healthcare9080947

Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. J Comput-Mediated Commun 3(2). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

MetaHuman Creator (2022) Available at: https://metahuman.unrealengine.com/

Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst E77-D(12):1321–1329

Norman D (2013) The design of everyday things. Basic Books, New York

Reeves TC (2006) Design research from a technology perspective. In: Van den Akker J, Gravemeijer K, McKenney S, Nieveen N (eds) Educational design research. Routledge, London, pp 52–66

Simon HA (1996) The sciences of the artificial, 3rd edn. MIT Press, Cambridge, MA

Slater M, Usoh M (1993) Representations systems, perceptual position, and presence in immersive virtual environments. Presence 2(3):221–223

The Design-Based Research Collective (2003) Design-based research: an emerging paradigm for educational inquiry. Educ Res 32(1):5–8

Unreal Engine 5 (2022) Available at: https://www.unrealengine.com/

Vaughn J, Lister M, Shaw R (2016) Piloting augmented reality technology to enhance realism in clinical simulation. Comput Inf Nurs 34(9):402–405. https://doi.org/10.1097/CIN.0000000000000251

Wagner I et al (2009) On the role of presence in mixed reality. Presence: Teleoperaters Virtual Environ 18(4):249–276. https://doi.org/10.1162/pres.18.4.249

Wang F, Hannafin MJ (2005) Design-based research and technology-enhanced learning environments. Educ Tech Res Dev 53(4):5–24

Weber S, Weibel D, Mast FW (2021) How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism (opinion). Front Psychol 12:1–10. https://doi.org/10.3389/fpsyg.2021.628298

Chapter 12
AI-Powered and "Augmented" Dentistry: Applications, Implications and Limitations

Rasa Mladenovic

Abstract The chapter deals with the possibilities that artificial intelligence (AI) and augmented reality (AR) provide in everyday dental practice and education, as well as the possibilities for the advancement of this technology in the future. The implementation of augmented reality and virtual reality in AI-powered dental technologies are rapidly evolving into applicable solutions for clinical practice and bridging the gap between the digital and physical world. The development of AR is leading to the development of a new field in dentistry, which can be called "augmented dentistry" that allows the use of new dental technologies to improve various aspects of clinical practice, virtual treatment planning, and the evaluation and modification of virtual treatment outcomes.

12.1 Key Concepts of Artificial Intelligence

Artificial Intelligence (AI) is a field of computer science that aims to perform specific tasks that require human intelligence. It allows to automatically extract important features of input data and discover previously hidden patterns. The development of AI systems has gained momentum in many areas of biomedicine and its use has become widespread in healthcare. In dentistry, AI can create a specific algorithm that further helps in diagnosis and treatment planning. In addition to diagnostics, it has found application in dental education, and it is also widely used in dental laboratories (Mladenovic et al. 2022c).

Neural Networks (NN) are information processing systems inspired by the analytical processes of the human brain, which are widely used to solve complex problems in the real world (Basheer and Hajmeer 2000). Neural networks can be significant in diagnosis and treatment planning, especially in diseases and conditions with a multifactorial cause or without a precise etiology. AI can also be applied to the

R. Mladenovic (✉)
Department for Dentistry, Faculty of Medical Sciences, University of Kragujevac, Kragujevac, Serbia
e-mail: rasa.mladenovic@med.pr.ac.rs

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_12

screening and classification of suspected altered mucosa, where it can detect single-pixel changes that are not visible to the naked eye. Artificial intelligence can also be used to predict genetic predisposition to oral cancer (Majumdar et al. 2018).

Machine learning is a field of artificial intelligence that has been applied for many years, primarily for data analysis. Machine learning is defined as a process that enables computers to learn without explicit programming. Today, machine learning has enabled a computer to classify or predict an outcome from a large database (Narang 2022). A variety of dental applications can benefit from machine learning models, including disease diagnosis, prognosis, treatment, and clinical workflow automation. In addition, machine learning has great potential to transform traditional healthcare delivery (Jung and Kim 2016; Gajic et al. 2021). A subfield of machine learning is deep learning, which allows machines to mimic human intelligence in increasingly independent and sophisticated ways, and then use the extracted knowledge to produce the best predictive results.

12.2 Basic Concepts of Medical Informatics

The development of digital dentistry has made it important to understand the basic concepts underlying medical informatics. In order to do this, it is helpful to know their origins.

DICOM (Digital Imaging and Communications in Medicine) is an open standard for handling, storing, printing and transmitting information in imaging. DICOM establishes a set of rules that allow integration of medical imaging devices from multiple manufacturers (Grant 2018). Today, manufacturers of almost all imaging systems implement the DICOM standard in their products. The DICOM format ensures that all data stays together in one file. The advantages of digital radiography technology are numerous, and one of the key ones is the ability to transmit digital data sets over a network. How fast the data will be transferred depends solely on the characteristics of the network. As human resources have to be involved in the transport of conventional films, image transmission via a data network from the imaging system to the diagnostician is an ideal solution. Also, conventional films require large archive systems that take up significant space, while digital archives require only a fraction of the space. Undoubtedly, the primary benefit for the staff comes from the fact that the images in the digital archive are sorted alphanumerically and are easy to search and can be downloaded at any time. Also, web technologies make it possible to access images and findings on any computer, in any location, with appropriate rights (security and confidentiality), using the Internet, and to write a radiological report (Schulze and Hoffmann 2011). The STL (Standard Triangulated Language) format describes the surfaces of bodies using triangles. The density of triangles depends on the initial resolution and mathematical algorithms (see Fig. 12.1). A DICOM file tends to provide more information about what is inside the volume, while an STL file tends to provide more information about the surface of the volume (Vecsei et al. 2021).

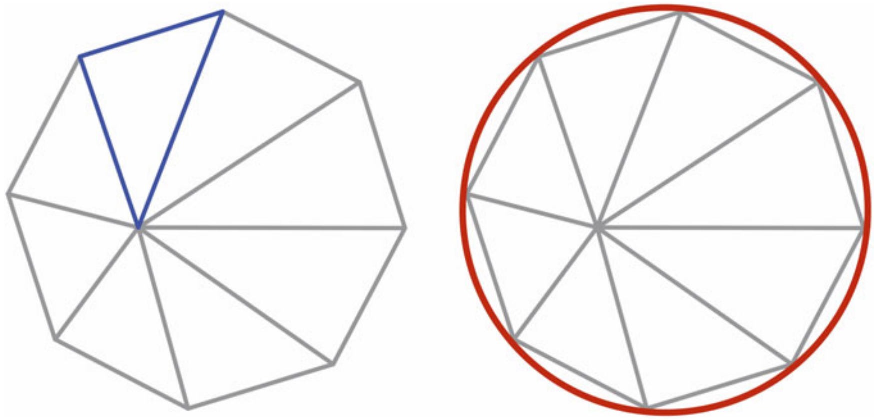


Fig. 12.1 STL model (left), CAD model (right). Image by the author

In digital streams, STL files representing oral structures can be obtained from multiple sources. The easiest and fastest way to create STL data is to use an intraoral scanner. In an indirect CAD/CAM (Computer-aided design and Computer-aided manufacturing) procedure, we can obtain STL data by digitalizing a plaster model based on an impression, using a laboratory scanner. At the same time, the DICOM data obtained from CBCT (Cone Beam Computed Tomography) devices can also be converted into STL data. There are many commercial (fee-based) and open source (free) software packages for exporting DICOM images to STL data (Kinariwala et al. 2021). The shape of the created STL model varies slightly from one software package to another. If STL files come from multiple sources, they can be aligned together with overlays and used to model or produce a number of modern design processes and tools for more precise and less invasive treatment. These advantages are used by navigational implantology, as well as guided endodontics. Three-dimensional (3D) printing has become easier with the advancement of technologies, as well as the development and reduction of hardware and software costs. Patient-specific 3D models are now used in many situations in oral and maxillofacial surgery, including education, surgical planning, and simulation. Of the approximately 100 3D CAD data file formats used as 3D source files in digital dentistry, the STL file format has found the most use as a 3D printing format. The quality of STL data affects 3D printing, and incorrect STL data can lead to unsuccessful 3D model creation (Kamio et al. 2020). In addition to 3D printing, the STL obtained by segmentation is of great use in creating ready-made 3D models for implementation in Augmented Reality (AR). Such models shorten the time necessary for modelling, and faithful copies of certain anatomical details are obtained.

12.3 AI-Powered and "Augmented" Dentistry

Virtual component of AI, known as software-type algorithms, is the main component used in dentistry. Because of their powerful capabilities in data analysis, these virtual algorithms are expected to improve the accuracy and efficacy of dental diagnosis, provide visualized anatomic guidance for treatment, simulate and evaluate prospective results, and project the occurrence and prognosis of oral diseases. AI has found its application in the fields of dental radiology, periodontology, endodontics, orthodontics, restorative dentistry and oral pathology (Narang 2022).

12.3.1 Artificial Intelligence in Dental Radiography

In order to simplify dental image analysis and thus speed up the treatment planning process, computerized assisted systems are being developed, such as Diagnocat, Overjet, Denti. AI, DentalXrai Pro. Diagnocat (Diagnocat Inc., USA) is an artificial intelligence system for dental diagnostics based on convolutional neural networks (CNN). Diagnocat AI system exploits a set of pre-trained semantic segmentation networks based on internally modified fully convolutional 3D U-Net architecture from to obtain voxel-perfect segmentation (Kurt Bayrakdar et al. 2021). Image analysis takes only a few minutes (around 2 min for OPG, and 5 min for CBCT), and the software generates a detailed report on the condition of each tooth with suggestions for further diagnosis (see Fig. 12.2).

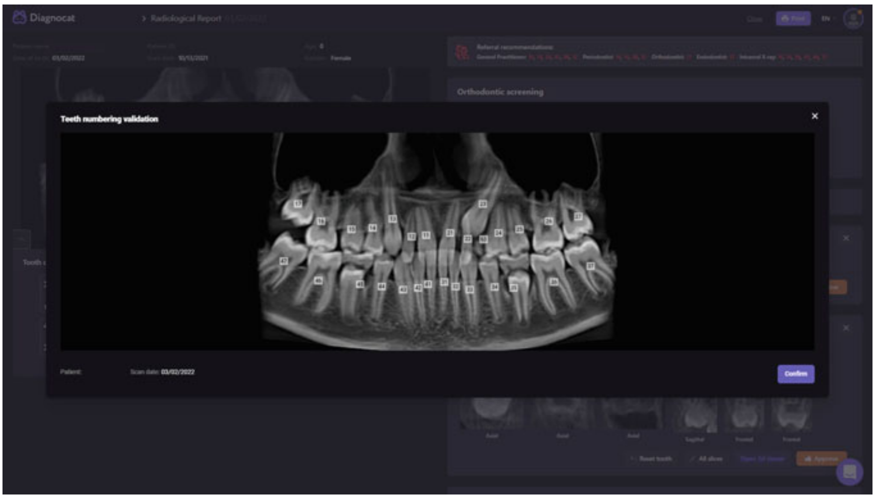


Fig. 12.2 Automatic tooth numbering. Image by the author

In addition to orthopan and simple images, Diagnocat enables the processing of more complex images, such as CBCT, and its special importance is reflected in the recognition of over 65 common dental conditions (see Figs. 12.3 and 12.4).

As many dentists do not receive extensive training on CBCT evaluation, and therefore are not competent enough for its analysis, there is a need for auxiliary learning, which greatly contributes to the development of such systems. Conventional tools for dental image analysis are quite demanding and pose a problem to dental

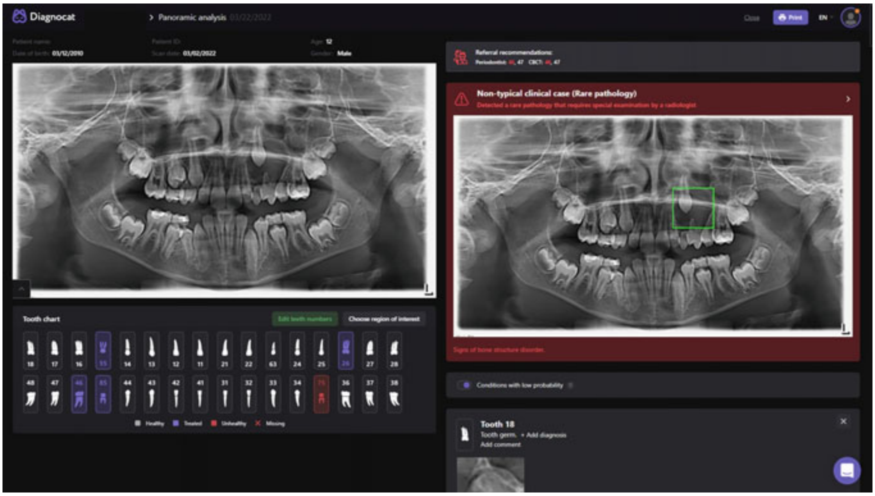


Fig. 12.3 Recognition of oral pathology. Image by the author

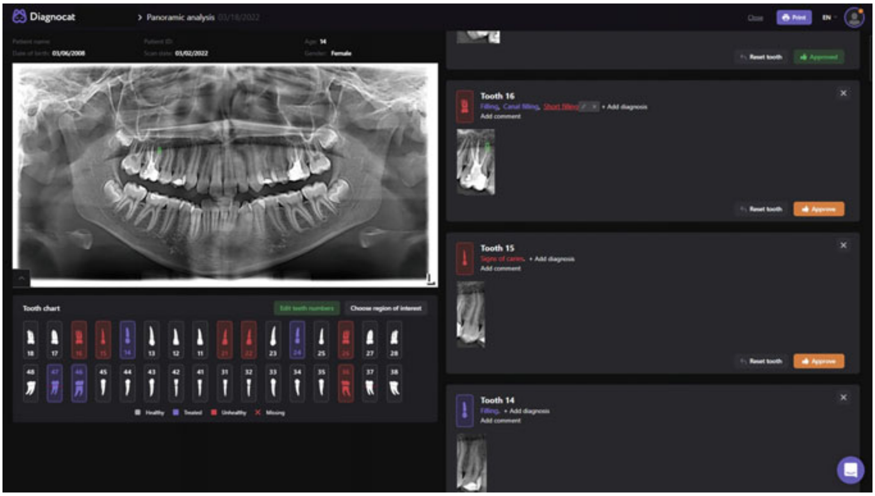


Fig. 12.4 Detailed report for each tooth. Image by the author

students due to limited availability and the complexity of interface, thus the advantage of AI systems is primarily reflected in their functionality, namely: ease of use, speed and accuracy (see Fig. 12.5).

The process of segmenting anatomical structures and converting DICOM format into STL files is significantly simplified. This facilitates 3D analysis and formulation of treatment plan, opening up numerous possibilities of digital dentistry based on STL (see Fig. 12.6).

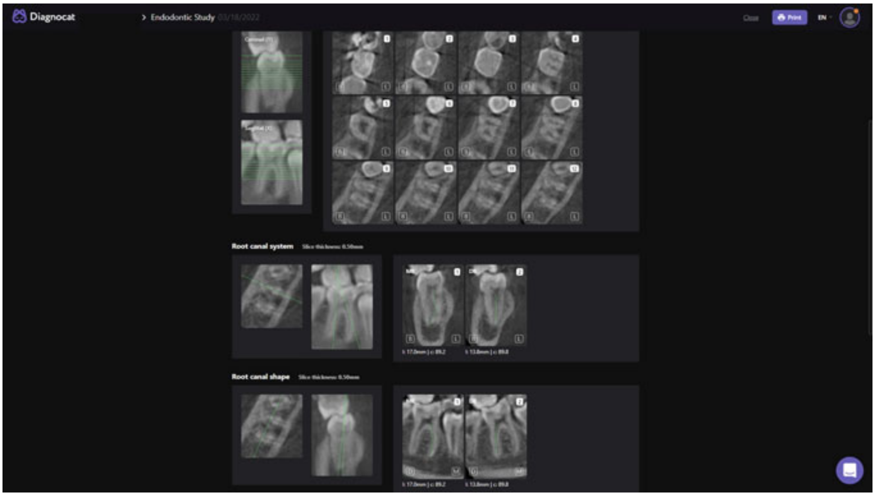


Fig. 12.5 Endodontic treatment planning section. Image by the author

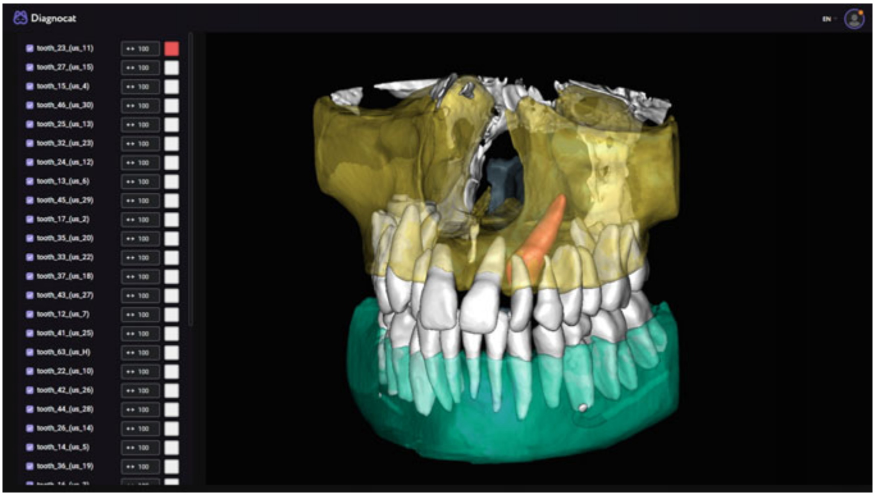


Fig. 12.6 STL viewer and automatic segmentation of oral structures. Image by the author

Using AI systems in radiographic interpretation may prevent wrong diagnosis and treatment planning as well as the unnecessary waste of time necessary for conventional analysis of CBCT images. Although AI systems have the potential to be applied in daily routine dental practice, they will not replace clinicians to a large extent, but will speed up and improve the diagnostic procedure itself. However, despite numerous benefits, this platform has some limitations. The detailed report Diagnocat generates does not provide information on redundant teeth which must be entered manually. Also, the selection dropdown does not include some of the most common paediatric dental issues. These are shortcomings that can be easily solved by updating the platform (Mladenovic et al. 2022c).

12.3.2 Artificial Intelligence in VR Systems

The expansion of 3D technology has led to the development of new resources and learning methods in dental education as well. These technological possibilities provide easy access to information, reuse of learning materials, practice and reinforcement of procedures or techniques. In addition to learning at one's own pace, this learning method is supported by the fact that the human brain functions on the principle of images and associations, that is, through visual experience.

In recent years, technological advances have made it possible to incorporate Virtual Reality (VR) simulation technology into preclinical dental education. Virtual Reality simulators provide the ability to integrate clinical scenarios into the operating environment and facilitate tactile skills using haptic technology. Haptics is a tactile feedback technology that takes advantage of user's sense of touch by applying force, vibrations or motion to the user (Mladenovic 2020a). Simodont (Nissin, Netherland) virtual simulator has been used for many years, and recently a new system called Virteasy Dental Haptic Simulator (HRV, France) powered by Unreal Engine (Epic Games, Inc. USA) has been introduced. Virteasy® haptic simulator consists of a PC-type computer, a touch-screen control for interacting with software, a pair of 3D stereoscopic glasses (Estar America® ESG6100), a plastic contra-angled handpiece connected to haptic device (Geomagic® Touch™ Ks Haptic Device) and a foot pedal to start the virtual handpiece in the simulator (see Fig. 12.7).

As part of the simulator software, there is an Assistant that has a rich library of procedures, as well as an Editor for data import, definition of therapeutic procedures and complete control of specific cases so that learning is at a higher level (see Fig. 12.8). The implantology module allows the user to import DICOM, add STL segmentation, and even select and plan the implant.

The Virteasy team is currently working on using AI for the improvement of the haptic rendering (the realism of virtual simulation compared to real clinical sensory feedback) by working with a big amount of data through the recording of professional dentist's gestures from surgeries to help guide students in their learning of motor skills during their initial pre-clinical training through the haptic motors.

Fig. 12.7 Virteasy V2/VR haptic simulator. Images by the author

12.3.3 Towards Implementing AI in AR

Augmented Reality (AR) is a simulation of a three-dimensional environment created using hardware and software that provides the user with realistic experiences and ability to interact. The application of artificial intelligence enhances the AR experience by enabling deep neural networks to replace traditional computer vision approaches and add new functions such as object detection, text analysis, and scene annotation. AR software used to use traditional SLAM (Simultaneous Localization and Mapping) computer vision techniques, but today AR applications rely on a deep learning model to provide more advanced features. AR developers can leverage AI algorithms to offer AR features such as enhanced interaction with the surrounding physical environment. Object labeling utilizes machine learning classification models. When a camera frame is run through the AR tracker, it matches the

Fig. 12.8 Virteasy assistant offers a rich library as well as the possibility of monitoring the educational process (up), Vireasy Editor allows importing patient recordings and creating unique educational content (down). Images by the author

image with a pre-defined label in the user’s classification library, and the label overlays the physical object in the AR environment.[1] The concepts of Augmented Reality and Virtual Reality are frequently confused with one another due to their similarities in their names and what they do. The difference is primarily in the technology used. Virtual Reality attempts to create an artificial world that a person can experience and explore interactively, through his or her senses, whereas Augmented Reality also

[1] 13 + AI Applications & Use Cases in Augmented Reality in 2022. https://research.aimultiple.com/ar-ai/.

brings about an interactive experience, but aims to supplement the real world, rather than creating an entirely artificial environment. Owing to intensive development of new devices for AR and VR, starting with Oculus Rift in 2012, and the launch of Google Cardboard in 2014, the application of VR in education and training came into focus of education research. However, a challenge both for the clinician and the patient is the usability of VR systems because the use of Head Mounted Displays (HMD) is problematic for some patients. As technology advances, new devices are being developed that offer hands-free transparent head-mounted displays (HMDs). One of them is the Microsoft HoloLens, that contains an inertial measurement unit, one front-facing depth camera and four side grayscale cameras used to map the spatial surrounding. As such, it is possible to display computer models and keep these holograms fixed, up to a certain accuracy, with respect to the physical space (Mladenovic and Djordjevic 2021; Frantz et al 2018). In dentistry, AR applications have found usage in the form of ready-made or dedicated applications, and increasingly high-quality graphics and details, ease of use and easy access allow for the daily development of new applications and learning models.

12.3.4 AR in Dental Education

Dental Simulator (Campinas, Brazil) is a mobile application for 3D simulation of local anaesthesia. The Dental Simulator application has three modes: study mode, simulation and augmented reality mode. The Simulation mode allows for simulation of local anesthesia in a 3D environment. After completion of each exercise feedback is provided on whether it was a success or there were errors, so users can reflect on their mistakes. Vuforia Engine-powered application (PTC, Parametric Technology Corporation, Boston, USA) brings immersive experience using camera of a mobile device or VR glasses. Image targets are the basis for image recognition technology (it is necessary to print a virtual patient target on A4 paper, as well as to stick a target on a 20 ml syringe). All student exercises are recorded using the University Mode. Through this mode, the educator is able to assess the learning process. The main challenge of simulation in an AR environment is image target detection (it requires well-lighted space and gentle manipulation). For improved detection of image targets, it is necessary to upgrade detection system and apply additional detection sensors. Also, it would be of major importance to utilize haptic technology to a greater extent. Further development should go in the direction of more realistic details, by adding involuntary movements of the patient, as well as the option of choosing the patient's age. This is especially important in managing inferior alveolar nerve block in children, where the position of the mandibular foramen differs significantly from that of adults (Mladenovic et al. 2019, 2020b, 2021, 2022a).

Immersify Dental (Manchester, UK) is a learning platform that combines 3D/AR technology and expert knowledge to create a unique learning resource. Through the application, users are able to explore the morphology of the teeth at their own pace.

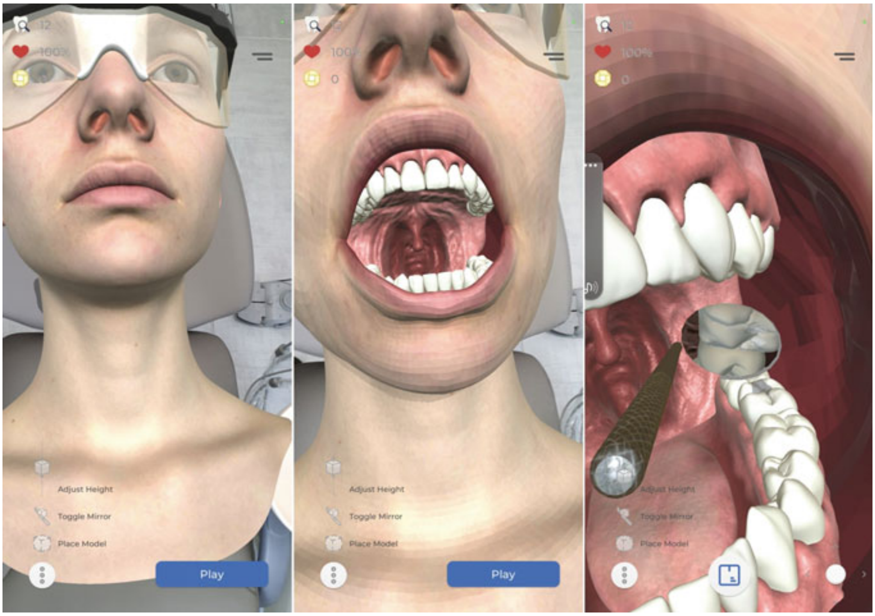


Fig. 12.9 Practicing a dental examination in an AR environment. Images by the author

Chartistry mode brings the experience of examining a virtual patient- students use a virtual dental mirror to identify teeth issues and record the data in a digital dental chart (Zsigmondy-Palmer, Universal, FDI) (see Fig. 12.9).

In addition to learning the dental examination and morphology of the teeth, it is also possible to learn the anatomy of the face and jaws through AR technology (see Fig. 12.10). The main advantage of this AR simulation is that it is delivered through a mobile device (iOS or Android). Also, the application offers the possibility of adjusting the virtual patient’s headrest, which can positively affect the acquisition of healthy postural habits during clinical examination (ergonomics). However, running the application causes intensive power usage as well as heating up of the mobile device, but these problems are of technical nature and are likely to be solved with the development in smartphone industry. It would be extremely important to create a virtual patient with primary and mixed dentition. In this way, students would be able to identify the type of dentition and understand the differences between primary and permanent teeth. To make the oral examination even more realistic, it is necessary to work on the implementation of haptic technology (Mladenovic et al. 2022b).

OMFA3D (3D Database of Clinical Oral and Maxillofacial Anatomy https://3dmodels.trevorthang.com) is a database of clinical maxillofacial anatomy developed in response to the lack of open access resources for anatomical education and the limitations of traditional atlases (Fig. 12.11). The database consists of over 100 3D models derived from CBCT scans and within each model there is an interactive 3D

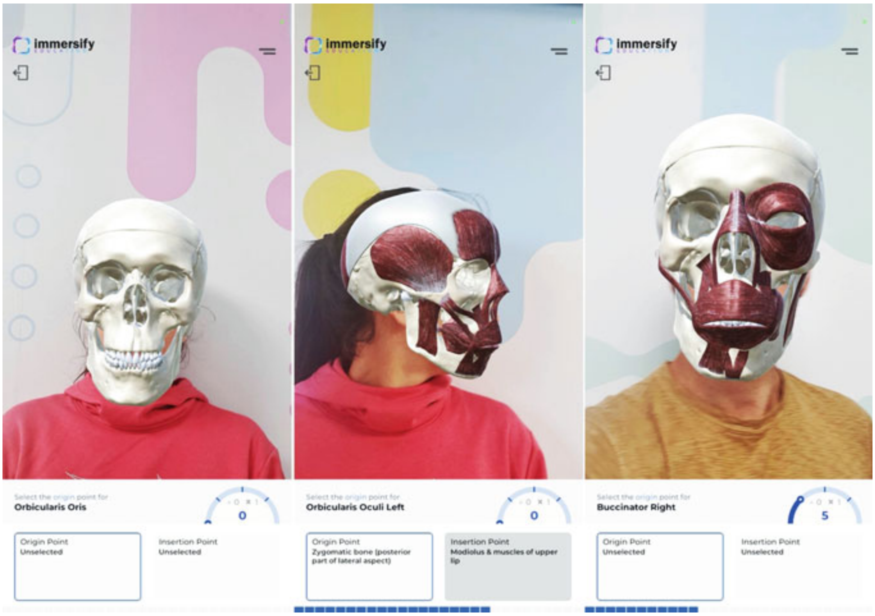


Fig. 12.10 Learning the anatomy of the face and jaws. Images by the author

viewer with anatomical markings, a description of the model and possible variations. For even greater interaction, it is possible to preview the model within the AR environment (Zhou and Thang 2022).

12.3.5 AR in Dental Practice

Despite the great need, only a few studies have dealt with the implementation of the AR-guided dental implant placement. In a review of two clinical cases reported by Pellegrino et al. (2019), AR was applied in implant placement planning. Prior to the procedure itself, the position of the implant was virtually planned, which contributed to a dynamic navigation system that was displayed on AR glasses. This, in turn, allowed for the use of a computer-aided procedure to occur. Ma et al. (2018). also proposed an AR navigation system for dental implant surgery. Their navigation system provides the visual field of the surgical site and the display of the overlaid scene at different depths. The positive results of these studies indicate that the AR-guided method of implant placement may prove to be an exceptional resource in implantology due to the simplicity of the procedures and the reduction of operating time. As such, it shows great potential in the future of implantology.

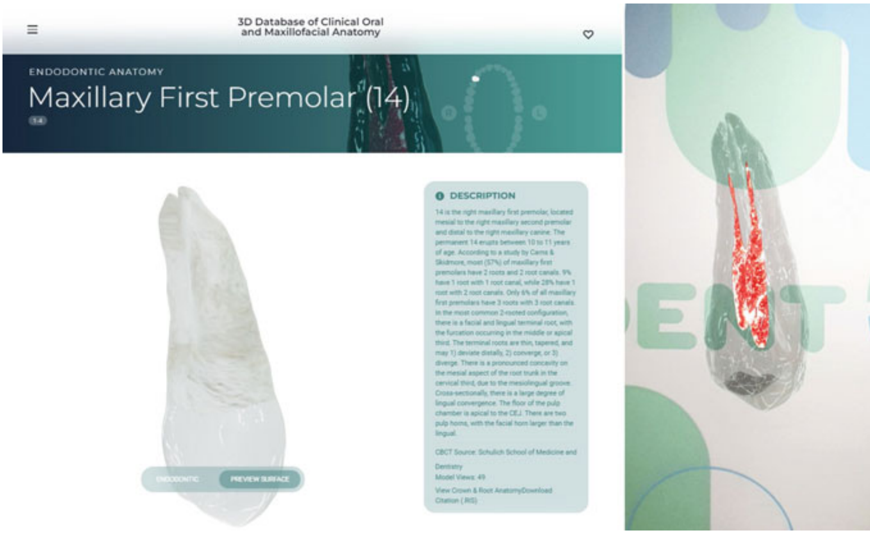


Fig. 12.11 Tooth morphology atlas (left), and view in augmented scene (right). Images by the author

12.3.6 AR Application for Dental Patients

In cosmetic dentistry, achieving the optimal appearance of planned restorations for the patient can be considered an important goal. A pre-visualization can be achieved with a conventional laboratory-made wax and intraoral model, in the form of a two-dimensional smile design by overlapping idealized teeth forms onto a portrait picture of the patient (Coachman et al. 2020). Recently, a procedure using a three-dimensional facial scan of the patient was proposed, but this requires a lot of time and the ability to combine different software, which is not yet simple for all the prosthetic laboratories (Stanley et al. 2018). To overcome this problem, a solution using AR software (IvoSmile, Ivoclar Vivadent, Schaan, Liecmobilhtenstein) was proposed, offering fast 3D conception. Before starting the rehabilitation of the aesthetic zone, the patient is able to see his future teeth through AR software, and this model represents a discussion platform where the patient can evaluate and visualize his opinion regarding the future position, length and shape of his teeth. The AR software allows for projecting different teeth length, width, and color onto the patient's face in real time (Touati et al. 2019). The patients can modify the smile projection until they feel comfortable with the result (see Fig. 12.12), and the selected AR projection can be saved and later imported into the CAD/CAM software (Touati et al. 2021).

Fig. 12.12 Co-diagnostic session with the help of augmented reality. Images reproduced under creative commons—Touati et al. (2021)

12.3.7 Serious Games

Today, more and more modern education models are based on serious games. Serious games, not games for entertainment purposes, are designed to simulate a real situation, allowing users to gain experience in a safe learning environment, without harming patients. Moreover, any mistake in the game can increase the student awareness in clinical practice, and the player loses the feeling of effort and repetition. Entertainment features of video games can also be incorporated to improve engagement and motivation for learning, including an appealing interface, audio integration, and high-quality graphic (Zafar et al. 2022).

12.4 Conclusion and Future Development

The progress of digitalization in dentistry has opened up new fields for the implementation and development of systems based on artificial intelligence, and deep learning models are already helping clinicians speed up the diagnosis, thus enabling better and faster treatment. When it comes to dental education, the use of VR and AR systems has been around for a long time. Their advantage is the acquisition of the necessary skills, achieved by redundant training and feedback. All of the above is very important for targeted guidance of students and ensuring a good teaching process (Zhang et al. 2022). The development of new devices of smaller dimensions will enable the application of VR and AR technologies in wider clinical practice, in areas where ergonomics and dimensions were the biggest problem. Future studies should focus on the formation of technological standards with high-quality data and the development of scientifically proven AR/VR gadgets for dental practice (Fahim et al. 2022). The progress of AI will bring new achievements, thus providing an opportunity for dental science to bring more innovations to the treatment and care of patients, as well as to facilitate the clinician's daily work. It is up to dentists to collect authentic data in order to get accurate results from the future AI model.

References

Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 43:3–31

Coachman C, Georg R, Bohner L, Rigo L, Sesma N (2020) Chairside 3D digital design and trial restoration workflow. J Prosthet Dent 124(5):514–520

Fahim S, Maqsood A, Das G et al (2022) Augmented reality and virtual reality in dentistry: highlights from the current research. Appl Sci 12(8):3719. https://doi.org/10.3390/app12083719

Frantz T, Jansen B, Duerinck J, Vandemeulebroucke J (2018) Augmenting Microsoft's HoloLens with vuforia tracking for neuronavigation. Healthc Technol Lett 5:221–225. https://doi.org/10.1049/htl.2018.5079

Gajic M, Vojinovic J, Kalevski K, Pavlovic M, Kolak V, Vukovic B, Mladenovic R, Aleksic E (2021) Analysis of the impact of oral health on adolescent quality of life using standard statistical methods and artificial intelligence algorithms. Children 8(12):1156. https://doi.org/10.3390/children8121156

Grant GT (2018) Craniomaxillofacial reconstruction based on 3D modeling. In: Greenberg A (ed) Digital technologies in craniomaxillofacial surgery. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-1532-3_4

Jung SK, Kim TW (2016) New approach for the diagnosis of extractions with neural network machine learning. Am J Orthod Dentofacial Orthop 149(1):127–133

Kamio T, Suzuki M, Asaumi R et al (2020) DICOM segmentation and STL creation for 3D printing: a process and software package comparison for osseous anatomy. 3D Print Med 6:17. https://doi.org/10.1186/s41205-020-00069-2

Kinariwala N, Buchgreitz J, Bjørndal L et al (2021) Endodontic Guides and Software Planning. In: Kinariwala N, Samaranayake L (eds) Guided endodontics. Springer, Cham. https://doi.org/10.1007/978-3-030-55281-7_4

Kurt Bayrakdar S, Orhan K, Bayrakdar IS et al (2021) A deep learning approach for dental implant planning in cone-beam computed tomography images. BMC Med Imaging 21:86. https://doi.org/10.1186/s12880-021-00618-z

Ma L, Jiang W, Zhang B et al (2018) Augmented reality surgical navigation with accurate CBCTpatient registration for dental implant placement. Med Biol Eng Compu 57(1):47–57

Majumdar B, Saroda SC, Saroda GS, Patil S (2018) Technology: artificial intelligence. BDJ 224:916

Mladenovic R, Djordjevic F (2021) Effectiveness of virtual reality as a distraction on anxiety and pain during impacted mandibular third molar surgery under local anesthesia. J Stomatol Oral Maxillofac Surg 122(4):e15–e20

Mladenovic R, Dakovic D, Pereira L et al (2020b) Effect of augmented reality simulation on administration of local anaesthesia in paediatric patients. Eur J Dent Educ 24:507–512. https://doi.org/10.1111/eje.12529

Mladenovic R, Mladenovic K, Milanovic P, Selakovic D (2021) Augmented reality technology as a method of distance learning for local anesthesia training. J Dent Educ 85(Suppl. 3):2038–2040. https://doi.org/10.1002/jdd.12581

Mladenovic R, AlQahtani S, Mladenovic K et al (2022a) Effectiveness of technology-enhanced teaching methods of undergraduate dental skills for local anaesthesia administration during COVID-19 era: students' perception. BMC Oral Health 22:40. https://doi.org/10.1186/s12903-022-02077-6

Mladenovic R, Matvijenko V, Subaric L, Mladenovic K (2022b) Augmented reality as e-learning tool for intraoral examination and dental charting during COVID-19 era. J Dent Educ 86(Suppl. 1):862–864. https://doi.org/10.1002/jdd.12780

Mladenovic R, Pereira L, Mladenovic K et al (2019) Effectiveness of Augmented Reality Mobile Simulator in Teaching Local Anesthesia of Inferior Alveolar Nerve Block. Journal of Dental Education 83:423–428 https://doi.org/10.21815/JDE.019.050

Mladenovic R, Milosavljevic M, Stanisic D, Vasovic M (2022c) Importance of artificial intelligence in the analysis of children's CBCT imaging by dental students. J Dent Educ 1–3. https://doi.org/10.1002/jdd.13060

Mladenovic R (2020a) The Usage of Augmented Reality in Dental Education. In: Geroimenko V (eds) Augmented Reality in Education. Springer Series on Cultural Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-42156-4_8

Narang D (2022) Artificial intelligence in dentistry. International Journal of Dental Science and Innovative Research 5(2):163–170

Pellegrino G, Mangano C, Mangano R et al (2019) Augmented reality for dental implantology: a pilot clinical report of two cases. BMC Oral Health 19:158

Schulze D, Hoffmann G (2011) Cone-Beam Computed Tomography and Navigation. In: Kramme R, Hoffmann KP, Pozos RS (eds) Springer Handbook of Medical Technology. Springer Handbooks. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74658-4_20

Stanley M, Paz AG, Miguel I et al (2018) Fully digital workflow, integrating dental scan, smile design and CAD-CAM: case report. BMC Oral Health 18:134. https://doi.org/10.1186/s12903-018-0597-0

Touati R, Fehmer V, Ducret M et al (2021) Augmented Reality in Esthetic Dentistry: a Case Report. Curr Oral Health Rep 8:23–28. https://doi.org/10.1007/s40496-021-00293-7

Touati R, Richert R, Millet C, Farges J-C, Sailer I, Ducret M (2019) Comparison of two innovative strategies using augmented reality for communication in aesthetic dentistry: a pilot study. J Healthc Eng 1–6

Vecsei B, Czigola A, Róth I et al (2021) Digital impression systems, CAD/CAM, and STL file. In: Kinariwala N, Samaranayake L (eds) Guided Endodontics. Springer, Cham. https://doi.org/10.1007/978-3-030-55281-7_3

Zafar S, Mladenovic K, AlQahtani S, Puranik C, Mladenovic R (2022) Assessing the pedological impact of local anesthesia dental simulator as serious game. Appl Sci 12(7):3285. https://doi.org/10.3390/app12073285

Zhang B, Li S, Gao S et al (2022) Virtual versus jaw simulation in oral implant education: a randomized controlled trial. BMC Med Educ 20:272. https://doi.org/10.1186/s12909-020-02152-y

Zhou KX, Thang T (2022) Rapid development of a novel and open-access mixed reality resource for dental education. J Dent Educ 86(Suppl. 1):783–786. https://doi.org/10.1002/jdd.12801

Chapter 13
Augmented Reality and Artificial Intelligence: Applications in Pharmacy

Don Roosan

Abstract This chapter addresses Artificial Intelligence (AI) and Augmented Reality (AR) applications in pharmacy and pharmaceutical sciences. AI collects and analyzes the data and mimics human cognitive functions to make intelligent decisions. AR allows a combination of real-world and computer-generated three-dimensional visualization that can be applied in education, training, and research. The development of AI and AR assists healthcare professionals in providing more personalized treatment options and planning for complex procedures to deliver better outcomes and quality of care. It can be utilized in pharmacy practice, including pharmacy education, pharmacogenomics, drug development, personalized medication, and treatment predictions. In addition, AR and AI can also help to create an interactive learning environment for pharmacy students to learn about real-world experiences from a classroom setting. Furthermore, the incorporation of AI and AR can help improve how healthcare providers utilize Electronic Health Records (EHR) to interact with patients efficiently to optimize medication outcome. With the help of these technologies, there is great potential for the future of pharmacy to be more advanced and efficient in delivering patient care.

13.1 AI and AR Application in Pharmacy Education

Computers use immense data and resources to provide us with personal comfort in many incredible ways. Technology has significantly changed our everyday lives, including our access to education and its delivery. The use of emerging technologies provides new opportunities for students in the pharmaceutical field to expand their clinical knowledge and communication skills. Artificial Intelligence (AI) and Augmented Reality (AR) can be utilized as separate teaching tools that allow students to engage in a learning environment mimicking real-world situations. AR provides

D. Roosan (✉)
Faculty of Pharmacy, Department of Pharmacy Practice and Administration, Western University of Health Sciences, Pomona, CA, USA
e-mail: droosan@westernu.edu

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_13

various data such as images, text information, and animations to create virtual simulations for different situations. It can be used in different areas of pharmacy, such as pharmaceutical product development, manufacturing, and drug discovery. AR teaching methods can be used in multiple courses in pharmacy school, such as medication counselling, medicinal chemistry, anatomy, Cardiopulmonary Resuscitation (CPR) training, and sterile compounding techniques training. For students who are visual learners, AR can be an excellent resource for educators when teaching essential concepts rather than trying to explain them verbally.

13.1.1 Medication Counselling

Pharmacists are licensed drug experts who are educated to be familiar with all medications on the market. Each medicine has its indication, specific form of administration, unique side effects, and drug-drug interactions. Medications, such as injections, inhalers, or nasal sprays, require additional consultation on the correct use and route of administration. However, it's not ideal for schools to provide every kind of medication on the market for students to view as this can be very costly and a great deal of medication waste. As a result, this may lead to a lot of uncertainty for students regarding counselling on those medications since they have never seen the products before. Nevertheless, AR can potentially be the solution to this problem. AR mimics real-world situations, so if a patient requires consultation on a new medication, the pharmacy student can take this time to practice working on their counselling as a pharmacist. The virtual images and videos can show the actual medication and devices inside the package, allowing students to have multiple attempts to practice learning the proper use of the medications. This learning tool would be beneficial in bridging the gaps between knowledge and practice by building the students' confidence, self-perceived competence, and communication skills.

Virtual Reality (VR) provides visual and sensory instructions and guides students to counsel the patients about medications (Fox and Felkey 2017). A study was done by AR MagicBook, using a web-based marker system called HP Reveal. HP Reveal is an AR platform that is easily accessible, user-friendly and can be implemented with any mobile device. A Magicbook is similar to a traditional book but contains AR images relevant to the study materials. In this study, naloxone was the medication of choice to analyze since the understanding of this drug therapy is vital to prevent death from opioid overdose, and a pharmacist's consultation on when to use it and how to administer a dose is significant. Students could access images of naloxone boxes, scripts, and videos to use for more guidance. A built-in AR marker allowed students to view these videos as 3D models, which were similar to real situations. As a result of the study, students quiz scores improved by approximately 42% compared to pre- and post-quiz grades, and students were thrilled with the new study methods. Thus, AR technology could be a powerful tool to help excel in a student's education.

Ultimately, AR technologies can help save costs instead of purchasing medications for hands-on learning for students (Schneider et al. 2020) and students can have access to these materials at their leisure.

13.1.2 Medicinal Pharmacy Chemistry

AR can also be used in medicinal chemistry, where students can make imaginary 3D structures of drug molecules. Medicinal Pharmacy Chemistry is an introductory course of the Doctor of Pharmacy curriculum, providing students with the fundamental knowledge and understanding of medications and their mechanism of action. Since textbooks and lecture slides show two-dimensional chemical structures, it can be challenging for most students to visualize the molecular structure in a three-dimensional image in their heads. Although there are molecular modelling kits available, it is time-consuming to make complex ones. In a study, computer graphics and AR were used to create study materials showing 3D drug molecules and their targets. The purpose of the study was to assist students with the visualization of 3D drug molecules and the feasibility of the AR-based lecture. The study material was easily accessed through any mobile device, and the students were more engaged in self-directed learning (Smith and Friel 2021). The software tool, Reality-Convert, allows you to convert the pictured chemical structures into 3D models and viewed by AR applications (Borrel and Fourches 2017). RealityConvert is very useful for small-molecule drugs and macromolecules like proteins. These AR technologies allow students to visualize complex drug molecules easily and enhance their learning retention. The result of the study shows that there was an improvement in teaching and learning medicinal chemistry based on students' feedback. The software tools allow students and researchers to easily visualize complex chemical structures as 3D molecules to understand medicinal chemistry better.

13.1.3 Anatomy

Human anatomy is also another required course for pharmacy students. Although technology has advanced significantly over the years, the methods used to teach anatomy have not changed. Most schools only provide a skeleton or human anatomical mannequin to observe during class, leaving students to use their textbooks after class. Researchers were interested in seeing if virtual reality could improve the way traditional anatomy is taught, so they conducted some studies implementing VR simulations of the human body in anatomy courses. Students were looking at a 3D human body and its dissections. A meta-analysis reviewed fifteen different randomized controlled trials. The studies compared traditional anatomy courses and a combination of VR technologies. The study's primary outcome was the VR intervention's on exam grades. With AR technologies, students can view 3D human bodies at any time using a mobile device. The results showed that the students using VR-based technology had significantly higher grades than the group of students using traditional study methods only. The secondary outcome was students' satisfaction levels with the addition of VR learning methods. The results showed that the students' satisfaction levels were higher with VR education and most students had a greater interest in learning with VR methods instead of the conventional way (Zhao et al. 2020).

In addition to using AR for academic settings, it can be used for training purposes. Pharmacists play an essential role in disease prevention and are well-positioned to facilitate vaccinations in the community. Immunization delivery is a required training within the curriculum for students in pharmacy school. Traditionally, students pair up with another partner and practice administering two intramuscular and one subcutaneous injection using normal saline. However, the proper vaccine administration technique is a skill that requires time and practice. Thus, student pharmacists must practice their vaccination skills on patients under the direct supervision of other pharmacists. Using AR technologies could prevent unnecessary patient harm due to lack of practice (Coyne et al. 2019). The virtual reality simulations allow students to practice their vaccination skills repeatedly and build self-confidence and competency. In a study, students were trained with mixed reality simulation techniques. Microsoft HoloLens head-mounted devices, GIGXR applications Holohuman and Holopatient were used for vaccination training for pharmacy students (Bushell et al. 2020). Students used Holohuman, an anatomy application that showed a holographic image of the human body and all the underlying structures to identify the muscle layers for the injection sites. Moreover, students went through scenarios where simulated patients experienced signs of anaphylaxis, which is impractical to learn in a classroom setting. This mixed reality training helps engage students intellectually by preparing them to have quick thinking skills during an unanticipated event and enhance their workforce skills with remote hands-on experience.

13.1.4 Cardiopulmonary Resuscitation Training

CPR is a lifesaving intervention that requires all healthcare professionals to be trained, including pharmacists and pharmacy students. The quality of the CPR performance is crucial as it directly affects blood circulation during emergency cardiac arrest. However, most healthcare professionals have problems retaining practical chest compression skills (Cheng et al. 2015). Even though all healthcare providers must renew their basic life support training every two years, it is hard for them to consistently perform correct and effective CPR in emergencies. It is rarer for pharmacists to perform CPR on patients due to the nature of their setting, as fewer emergencies occur in the pharmacy. Traditional CPR training is done by watching video instructions on properly performing CPR. Then, a group practices doing CPR on mannequins with one or two instructors supervising. With most conventional ways, learners usually do not receive much or if any feedback on their quality of CPR abilities. However, utilization of VR CPR training enables you to receive real-time scoring and skill evaluation. It creates a realistic experience that allows the trainee to witness someone suffering a medical emergency, tests their readiness to act when needed and aids in retaining the information learned. This would improve CPR training engagement with virtual mannequin simulators (Kuyt et al. 2021). A recent study focused on measuring the effectiveness of VR-based CPR training and the instructors' perception of VR-based CPR training. The study used Google Glasses and Microsoft HoloLens as the virtual mannequin simulators. The VR technologies provided tutorials and simulations for the users on proper CPR techniques. VR headsets and controllers were utilized to track the users' location and input of CPR. The VR technologies measure whether the users provided effective CPR and immediate feedback. The results show that VR and AR technologies could play a big role in the future and be utilized as CPR training tools. The instructors agree that VR CPR training would increase students' attention and motivation to learn. Thus, more advanced technologies would allow healthcare providers to enhance their CPR skills.

13.1.5 Sterile Compounding Techniques Training

AR and AI technologies are also utilized in hospital pharmaceutical technology training. Sterile compounding is important when preparing various intravenous (IV) medications for patients. Sterile compounding requires technicians to use aseptic techniques to prevent contaminations from pathogens. Each IV medication has its package inserts and additional instructions to be used for compounding. These medications are typically very expensive, and hospitals do not want to waste drugs, so technicians use sterile water to practice their skills. However, this still includes waste of supplies and learners do not get to practice multiple times due to the availability of qualified sterile compounding space. In a recent study, AR was utilized with virtual cleanrooms and 3D glasses for sterile compounding training with the intent of helping technicians practice sterile compounding techniques as accurately as possible. A systematic review provided an overview of current VR simulation in hospital pharmaceutical technology. The purpose of the study is to see the feasibility of VR simulation in the hospital setting. The VR shows a virtual cleanroom with a laminar flow hood and all necessities. This allows the technicians to practice multiple times before they compound the actual IV medications, and the VR system would detect contaminations (Garnier et al. 2021). In the the study, 3D glasses were used to depict a virtual clean room to teach students the essential United States Pharmacopeia (USP) 797 standards to follow. Then, students chose the proper fitting hood and IV fluids for the different medications in a virtual simulation. Next, they began preparing the virtual IV products using the aseptic techniques they learned to prevent cross-contamination (Patel et al. 2011). Ultimately, VR technologies should be utilized for pharmacy training because this allows students to practice IV compounding without wasting any medications.

13.2 Mobile App Development for Pharmacy

As life expectancy increases, many patients suffer from chronic diseases (Eggerth et al. 2020). Patients with chronic illnesses must manage their health daily to prevent complications or slow disease progression. In some cases, patients must take multiple medications at different times of the day, monitor their caloric intake, or measure their blood pressure or blood sugar level daily. However, many patients forget to follow these daily tasks, and this can negatively impact a patient's overall health. Moreover, we are facing a growing shortage of doctors and getting an appointment to see your primary care provider has become more complex. It has become a burden for patients to see their doctor often due to the cost, transportation barriers and long wait times. Both patients and doctors agree there is insufficient time to provide the best care. Therefore, some mobile applications are developed as healthcare tools to guide patients in managing their chronic diseases. AI and AR are utilized in developing these mobile applications to provide multiple functions such as reminding patients to take medications, guiding patients in the correct administration of their medications, providing dietary support, regular physical activity support, and providing patient education.

Maintaining good medication adherence is the most significant way to prevent complications and slow disease progressions. Patients must take medications as directed by the doctor for the best effectiveness. Fortunately, many smartphone applications can aid patients with medication adherence. Most new smartphones and medical devices have built-in near-field communication (NFC) capabilities that monitor and record patients' medication adherence. This technology can also be applied to medication blister pack cards and help patients record when they took their medications. Once the NFC tag on the medicine is activated, the smartphone with a close enough application will automatically record the medication intake time.

Furthermore, there is a new technology that uses sensors for medication bottles. The smartphone application will recognize that the bottle is open and keep track of the remaining pills inside. It will alert patients to refill their prescription when it is low (Eggerth et al. 2020). AI applications have access to smartphone cameras and use the computer vision algorithm to identify the patient, medication and confirm the administration of drugs. The study provided strong evidence to improve medication adherence in the patients monitored by the AI smartphone applications (Babel et al. 2021).

Error in medication self-administration (MSA) causes poor medication adherence. Even though the patient adheres to medications as scheduled, it can cause treatment failure without proper administration. For example, it usually happens with non-oral medicines such as insulin, inhalers, nasal sprays, or even eye drops which require correct techniques. Despite doctors and pharmacists providing consultations on administering the medication accordingly, patients often forget. If healthcare professionals spent more time with their patients and directly observed the patient's administration techniques, these errors could be prevented, but sadly this is not always possible. Therefore, a study was done to detect and monitor MSA errors using AI to analyze the wireless signals in the patient's home. The AI-based solution does not require cameras to see a patient's movements. Instead, it requires a Wireless Fidelity (Wi-Fi)-like sensor that transmits signals around Wi-Fi frequency range using frequency-modulated continuous-wave (FMCW) radar system that works similarly to radio sensors (Zhao et al. 2021). The sensor transmits low-power signals, and the system analyzes the reflections of the signs and modulates the patient's movement. Ultimately, this system detects the patient's medication administration skills. The study showed wireless sensor observes the patient's inhaler administration skills. The wireless sensor can follow if the patient shakes the inhaler device, exhales before administration, and inhales the dose correctly. Then, if an error in MSA is identified, the health professionals can access these records and provide patient education on how to administer a dose correctly. As a result of the study, the AI sensor could appropriately identify the patient's error in MSA and improve medication adherence.

It is important for patients to monitor their caloric intake and maintain their physical activities to manage diseases. A smartphone application called "FoodLog" helps keep track of and record the patients' daily intakes. If the patients take pictures of their meals and the application finds the foods' nutrition profile and portion size. Also, the application provides feedback on their nutritional intake so that they can adjust their diet. Another smartphone application called "GoCARB" identifies the meal's carbohydrate content with simple pictures of patients' meals. Also, patients can utilize AR and AI to aid them with their physical activities. Movement-based video games, such as Nintendo Wii, were used in this study. The Nintendo Wii Fit Plus provides fitness video games in a fun way to improve patients' physical activities and get them moving. The study results showed significant improvements in patient's physical activities compared to the control groups (Rollo et al. 2016).

Finally, healthcare providers should always use layperson's terms when counselling their patients or giving consultations. Most patients are unfamiliar with medical terminologies, and some are not fluent in English or have low health literacy. Overall, people with poor health literacy tend to have treatment failure due to the lack of understanding of their medications. Smartphone applications with AR and AI are being utilized to help patients quickly access consultations regarding their medications. The smartphone application that uses AR and AI offers more information about the medication simply by taking a picture of the packaging. The application will show the medications' indications, dosing schedule, and side effects, provide additional information about the medication, and include YouTube video links to help understand in more detail. The result of the study demonstrated improvement in patient literacy (Ahmadvand et al. 2018).

Nowadays, many smartphone applications are being utilized as healthcare tools. The advancement in technology has not only improved healthcare but also allowed healthcare professionals and patients to become more comfortable in managing their disease states together efficiently. These applications will help patients' adherence to their medications, decrease healthcare costs and improve health outcomes. Thus, it would lead to ordering the disorders successfully.

13.3 Drug Development

The history of medicine has rapidly progressed and changed enormously over the past few centuries. Life expectancy in the United States has more than doubled compared to 160 years ago (the *United States* n.d.). In the past, many people died from simple infections to severe diseases due to a lack of treatment and understanding of diseases. With the development of medical sciences, we have conquered diseases that spread quickly, such as measles and chickenpox. Moreover, cancer and Human Immunodeficiency Virus (HIV) are no longer considered the deadliest diseases. With many treatment options available, we can now target the cause and help control it. There is still much research to be done to conquer these diseases and hopefully, we can find a cure. The development of new technologies allows the targeted delivery of therapeutic agents to maximize effectiveness and minimize adverse effects. AI and AR are innovative technologies utilized in drug development to accelerate the process and time it takes to get a drug on the market and reduce costs. The technologies are also applied in pharmacogenomics and help with drug development by providing a quicker, more accurate and cost-effective approach that allows patients to have a personalized medication therapy based on their genotype and inherited gene. Researchers used machine learning algorithms to identify chemical compounds, drug targets, effectiveness, and toxicities. Furthermore, AR can provide 3D visualization of chemical compounds to identify better target sites and interactions between the drug molecules and targeted antibodies. Thus, the new technologies save time and money in developing effective drugs for various disease states (Gupta et al. 2021).

Nowadays, many drug manufacturers aim to develop specialty medications known as biological medications. Biological medications are complex molecules that involve living organisms to treat or slow down the progression of severe diseases, such as cancer, by targeting specific parts of the human immune system. There is a type of biologics that mimics human antibodies, known as monoclonal antibodies. Understanding the structure of antibodies and their target sites is necessary to develop monoclonal antibodies. The antibody has two binding sites; two antigen-binding sites and one antibody receptor protein binding site. To make therapeutic antibodies, researchers must predict what type of receptor or antigens would bind to the monoclonal antibodies to prevent unwanted side effects. However, this process is very time consuming and challenging with 2D imaging. Thus, researchers can overcome these obstacles by utilizing AI and AR to visualize the monoclonal antibodies. This will allow them to predict the effects of medications by looking at their target sites via virtual platforms. The study used a smartphone application to view the large protein molecular structures in 3D models. Protein structures were downloaded from online protein databanks, and these images were processed using PyMOL, UCSF Chimera, and Blender to create 3D protein models. The protein structures were differentiated with different colours to highlight their properties. This application allows you to visualize the protein structures as 3D models and how proteins interact with other small molecules. The application enables users to rotate, zoom in and out, and create animations of protein structures. Incorporating this new technology would save us time and money in developing new medications (Chan et al. 2020).

The purpose of taking medication is to help control or treat disease and illness. However, sometimes the medicine can cause unwanted or unpleasant effects. Thus, it is essential to consider the side effects and drug safety when developing new drugs. Drug safety profiles are observed during drug development and clinical trials. Sometimes there are no observations of severe side effects during the drug development processes, yet findings of unwanted side effects during clinical trials. These side effects could be due to many factors, such as patient diversity, drug-drug interactions, or off-target interactions. This can be a severe problem as it can cause harm to actual patients. Therefore, AI and machine learning are utilized in observing and improving drug safety. Deep learning, a subset of AI, uses immense data to build complex models to assess toxicity and safety drug profiles. These models allow us to examine drug toxicity studies in pre-clinical trials while considering different patient ethnic backgrounds and other pertinent factors. It also helps minimize drug toxicities during clinical trials to prevent unwanted patient effects. Lastly, this helps save time, reduce costs and recognize ethical problems that may occur in developing new drugs (Basile et al. 2019).

Researchers from another study showed that AI had been utilized in many processes of drug development, such as peptide synthesis to molecular design, virtual visualization of molecular docking, drug repositioning, and protein misfolding to protein–protein interactions (Gupta et al. 2021). AI can identify the target molecules for certain medications and validate the target, screen the molecules against the desired target, validate affinity, selectivity, metabolic stability, and oral bioavailability. Thus, AI helps save many steps in identifying suitable molecules that could lead to future potential medications. AI can also screen millions of chemical compounds to determine specific drug targets using high throughput screening strategies (Tripathi et al. 2021). Furthermore, there are strategies called high throughput virtual screening, which allows for visualization of molecular docking, pharmacophore modeling, and drug simulations.

AI and AR have significant potential to find more chemical molecules that could be used for future drug components. Technological advancement of AI and AR can expedite the drug development process. Furthermore, patients would experience fewer side effects from new drugs due to the AI-implemented pre-screening of drug toxicity profiles. In the future, severe diseases like cancer will be treated as curable diseases with the production of newly developed medications.

13.4 Pharmacogenomics Education

Variation in the genetic sequence is what makes each human very different and unique. However, traditional medicine treats all humans in the same way, even though the genetic components of each human are different. Therefore, some medications may work well for certain people, while some drugs can cause more side effects.

Pharmacogenomics is a field of research that seeks to understand the correlation between human health and genetic variation (Cecchin and Stocco 2020). It provides patients with personalized medication based on their genomic information, which leads to better efficacy and reduced adverse drug reactions. Even though pharmacogenomics benefits patients, providers and pharmacists, it is not utilized widely in clinical settings due to a lack of knowledge of pharmacogenomics and its clinical implications. Therefore, it is essential for student pharmacists to learn and understand pharmacogenomics for future clinical practice.

Understanding a challenging subject, such as pharmacogenomics, can be hard for some students when using the traditional education system that shows molecular structures as 2D pictures on lecture slides or in textbooks. To better understand genes and their variation, AR is utilized for the 3D visualization of molecular structures. HoloLens, the mixed reality smart glasses made by Microsoft, is the most commonly used AR tool. It allows users to visualize the technological virtual graphics that can be incorporated into real-world backgrounds. AR tools enable students and researchers to visualize molecular structures in 3D and help convey complex concepts for better understanding. Studies show that AR-implemented learning systems improve students' learning outcomes by keeping them engaged. The study aimed to develop and assess the AR application PGxKnow, using HoloLens with AI capabilities for a better educational delivery system (Roosan et al. 2022).

PGxKnow is the AR application for pharmacogenomic education, which was developed by integrating educational concepts, information, and 3D molecular structures into an AR tool known as HoloLens. PGxKnow has three interactive designs that help users operate the application. The three design functions are gaze, movement, and voice. The gaze function uses the head position and direction of the user to track the user and move the 3D object through the HoloLens. The movement function uses the user's finger movements to rotate, drag or zoom in and out the holographic objects. Lastly, the voice function allows users to command orders to the application to go to the following concepts, relocate things or restart.

With these valuable functionalities, PGxKnow works as an educational tool for students by providing several video modules and 3D molecular structures with textual information. It starts with the simple concepts of pharmacogenomics, as seen in Figs. 13.1 and 13.2, which shows the 3D holographic nucleated cell used in the PGxKnow application. The details of the cell can be seen with its organelles in a 3D image. On the left side of Fig. 13.1, multiple voice command orders can be used in the application, allowing the user to direct the application to show different tasks verbally. The right side of Fig. 13.1 offers the textual explanation of the 3D nucleated cell. The nucleus is packed with chromosomes and each chromosome is made of a long strand of tightly coiled DNA, as seen in Fig. 13.3. The picture displays a holographic view of the 3D chromosomes and DNA structures. Another example is shown in Figs. 13.4 and 13.5, which is the concept of DNA mutation. There are voice command instructions exhibited on the left side and textual information on the right side. Finally, Figs. 13.6 and 13.7 show the protein folds with mutated genes, where the green and red colours are the mutated protein parts. These proteins serve as drug-metabolizing enzymes that function differently. The AR application design delivers a different approach to improving the education of pharmacogenomic concepts. It also allows students to bridge the gap between the clinical application element and understanding the principles of genomic medicine.

Finally, PGxKnow assessed its usability using a system usability scale (SUS) survey. Usability is measured by achieving specific goals with effectiveness, efficiency, and satisfaction. For this study, fifteen participants completed the SUS survey and analysed the results using SAS software. The average SUS score for the fifteen participants was 83, which was above average. The high SUS scores prove that the

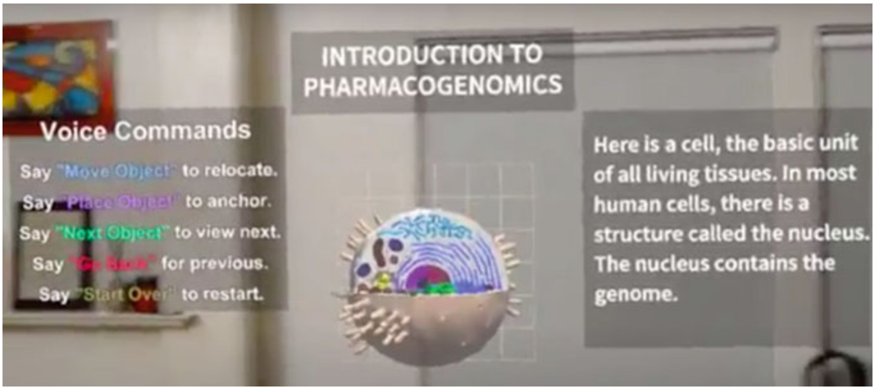


Fig. 13.1 The 3D model of nucleated cell structures. The hologram shows the nucleated cell structures with various organelles. On the left side, there are multiple voice commands. The right side offers the textual explanation of the 3D nucleated cell

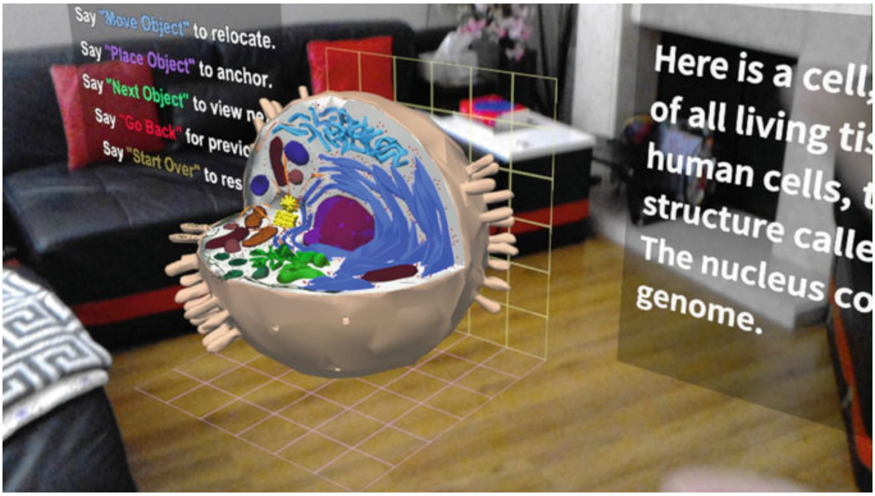


Fig. 13.2 Side view of the 3D model of Nucleated cell structures. This figure is the same as Fig. 13.1 but at a different angle

Fig. 13.3 The 3D model of DNA and chromosomes. This hologram is the side view of the 3D structures of the chromosomes and DNA. On the right side, it has the explanation of chromosomes in the nucleus, and each chromosome is made of a long strand of DNA

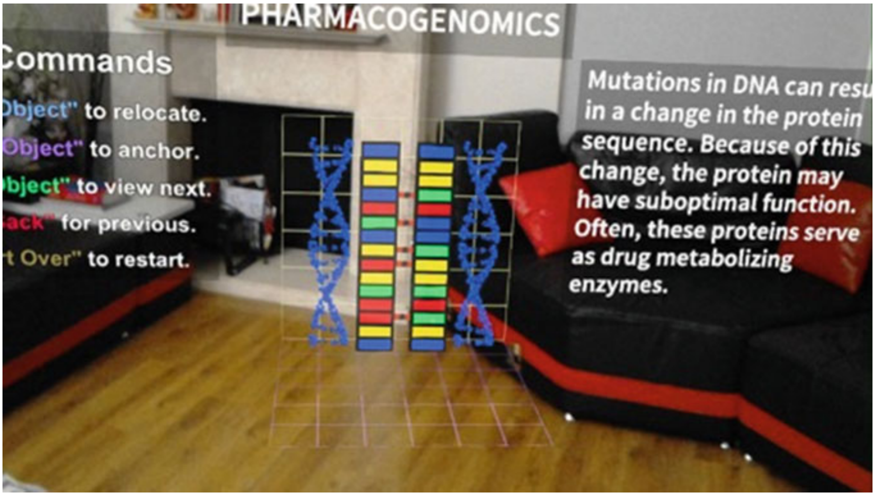


Fig. 13.4 The 3D model of mutations in DNA. On the left side, there are voice command instructions. On the right side, there is textual information to explain that the mutations in DNA can lead to different levels of functions for the enzymes that are made of proteins

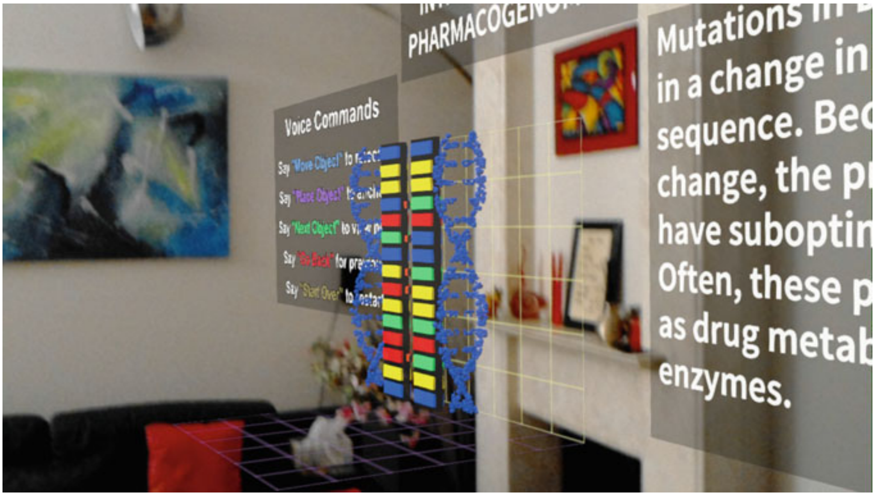


Fig. 13.5 Side view of the 3D model of mutations in DNA. This hologram is the same as Fig. 13.4, but from a different angle to show the 3D visualization

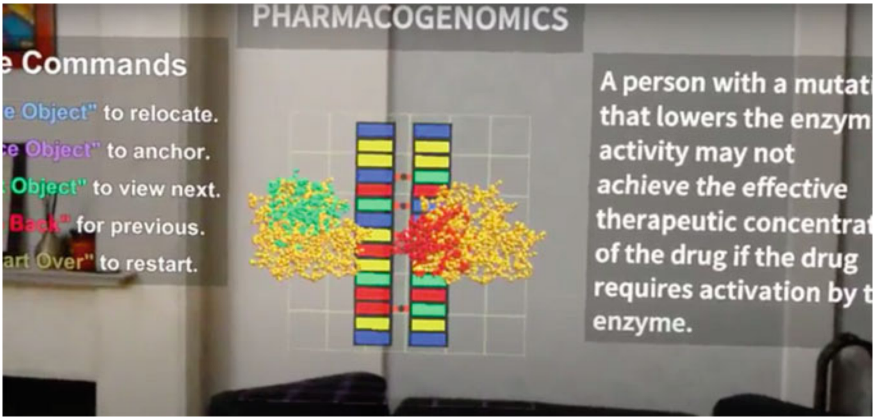


Fig. 13.6 The 3D model of protein folds with mutated genes. The green and red colours are the protein parts that are mutated. Due to the mutation, each enzyme responds differently to the medications that require activation of the enzyme

application is effective and efficient in teaching students and that the users' satisfaction level is high. PGxKnow demonstrates that the AR application, in conjunction with AI, can be utilized as an effective pharmacogenomics educational tool for everyone.

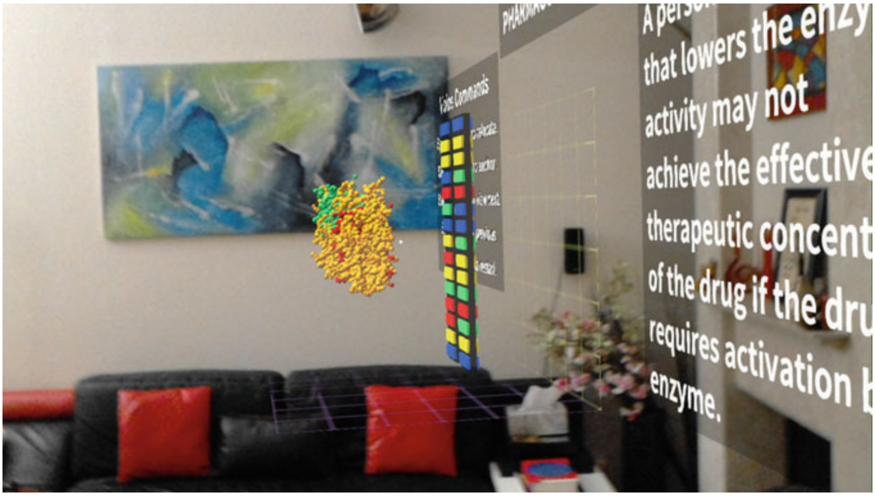


Fig. 13.7 Side view of the 3D model of protein folds with mutated genes. This hologram is the same as Fig. 13.6, but at a different angle to show the 3D visualization

13.5 Improve Clinical Care

"Treat the patient, not the number". It means that health professionals need to listen to their patient's medical complaints, symptoms, and overall situation. Healthcare professionals should not just focus on medical records or lab values to reach specific targeted goals. The current health information technology forces doctors to focus more on lab values and medical records than the patient. Wait times for patients to see their doctor are getting longer, and the actual visit is becoming more concise. Even during those short visits, doctors usually focus on documenting the patient's health status on the computer system instead of acknowledging their patients for making real connections. The electronic health records (EHR) system does not allow doctors to create authentic relationships with patients (René 2022). EHR engagement requires much attention since the doctor must review the patient chart, record patient demographics, and check each lab value for abnormalities. Also, doctors document patient complaints, health status, progressions, and treatment plans in the EHR per health insurance plans. As a result, doctors are not making much eye contact with their patients during visits but look at the computer screens instead. In the healthcare system, this is a significant setback from the development of technology. Patients want to feel connected with their doctors. Nevertheless, there is always room for improvement, and various functions. The next generation of design can be applied to EHR to enable physicians more time to interact with patients to improve overall care. New technologies can support doctors and patients using speech recognition to improve overall communication.

Nowadays, all healthcare systems have adopted EHR for better patient care, improved health outcomes, and reduced medical errors. EHR allows different health facilities to share patient information quickly and stores the record for a long time. However, during doctor's visits, physicians document what patients say about their health in EHR for insurance and record-keeping purposes. Doctors often focus on the computer to type information with minimal eye contact, leaving patients unsatisfied with the visit and the care. Patients desire to be the centre of the provider's attention. Unfortunately, most doctors do not have enough time to sit and listen to patients. However, using the speech recognition function in the EHR using AI, doctors can focus more on patients by having more eye contact. In a study, the researchers applied speech recognition technology in EHR systems with Morae usability software and compared it to the typing function. The speech recognition function had more extended notes, fewer typing errors and better quality than the typing function (Blackley et al. 2020). Another study showed significant improvements in physician satisfaction, documentation quality, and efficiency with the speech recognition function (Saxena et al. 2018).

In the Pre-EHR era, patients' medical records would pile up with tons of information for doctors to review. Unfortunately, doctors do not have time to read all the specific details and may even miss some information that could lead to substantial health problems for patients. EHR allows physicians to record a patient's health status electronically and refer to this information whenever needed. EHR can also summarize the essential health records for the doctors to review shortly before patients' appointments. An EHR typically consists of two types of data, structured and unstructured data. Structured data refers to the medications ordered, laboratory test results and procedures conducted. Unstructured data refers to progression notes, discharge summaries, and telephone encounters. To effectively summarize the patient's current status, it is essential to consider both data. A recent study extracted information from structured data with a timestamp and utilized with unstructured data to form a complete information set. Also, they used a Note-level Insights system to gain an overview of the patient's problem list and supported the physicians to become aware of the specific problems for a visit. As a result, the system could summarize more than 30,000 EHR updates daily and enabled physicians to reach the desired performance goals using AI (Suryanarayanan et al. 2021).

New technologies integrating AI can support both physicians and patients to improve overall healthcare. AI and AR integrated within the EHR, has the great potential to reduce healthcare waste and improve medication optimization. AI and AR have the potential to improve pharmacy education and various other practical healthcare applications.

References

Ahmadvand A, Drennan J, Burgess J, Clark M, Kavanagh D, Burns K, Howard S, Kelly F, Campbell C, Nissen L (2018) Novel augmented reality solution for improving health literacy around antihypertensives in people with type 2 diabetes mellitus: protocol of a technology evaluation study. BMJ Open 8(4):e019422. https://doi.org/10.1136/bmjopen-2017-019422

Babel A, Taneja R, Mondello Malvestiti F, Monaco A, Donde (2021) Artificial intelligence solutions to increase medication adherence in patients with non-communicable diseases. Front Dig Health 3:669869. https://doi.org/10.3389/fdgth.2021.669869

Basile AO, Yahi A, Tatonetti NP (2019) Artificial intelligence for drug toxicity and safety. Trends Pharmacol Sci 40(9):624–635. https://doi.org/10.1016/j.tips.2019.07.005

Blackley SV, Schubert VD, Goss FR, Al Assad, Garabedian PM, Zhou L (2020) Physician use of speech recognition versus typing in clinical documentation: a controlled observational study. Int J Med Inf, 141:104178. https://doi.org/10.1016/j.ijmedinf.2020.104178

Borrel A, Fourches D (2017) RealityConvert: a tool for preparing 3D models of biochemical structures for augmented and virtual reality. Bioinformatics (oxford, England) 33(23):3816–3818. https://doi.org/10.1093/bioinformatics/btx485

Bushell M, Frost J, Deeks L, Kosari S, Hussain Z, Naunton, M (2020) Evaluation of vaccination training in pharmacy curriculum: preparing students for workforce needs. Pharm: J Pharm Educ Pract 8(3):151. https://doi.org/10.3390/pharmacy8030151

Cecchin E, Stocco G (2020) Pharmacogenomics and personalized medicine. Genes 11(6):679. https://doi.org/10.3390/genes11060679

Chan K-F, Poh JJ, Wu WL, Gan SKE (2020) Augmented reality in scientific visualization and communications: a new dawn of looking at antibody interactions. Antibody Ther 3(3):221–226. https://doi.org/10.1093/abt/tbaa021

Cheng A, Brown LL, Duff JP, Davidson J, Overly F, Tofil NM, Peterson DT, White ML, Bhanji F, Bank I, Gottesman R, Adler M, Zhong J, Grant V, Grant DJ, Sudikoff SN, Marohn K, Charnovich A, Hunt EA, … for the International Network for Simulation-Based Pediatric Innovation R, Education (INSPIRE) CPR Investigators (2015) Improving cardiopulmonary resuscitation with a CPR feedback device and refresher simulations (CPR CARES study): a randomized clinical trial. JAMA Pediatr 169(2):137–144. https://doi.org/10.1001/jamapediatrics.2014.2616

Coyne L, Merritt TA, Parmentier BL, Sharpton RA, Takemoto JK (2019) The past, present, and future of virtual reality in pharmacy education. Am J Pharm Educ 83(3):7456. https://doi.org/10.5688/ajpe7456

Eggerth A, Hayn D, Schreier G (2020) Medication management needs information and communications technology-based approaches, including telehealth and artificial intelligence. Br J Clin Pharmacol 86(10):2000–2007. https://doi.org/10.1111/bcp.14045

Fox BI, Felkey BG (2017) Virtual reality and pharmacy: opportunities and challenges. Hosp Pharm 52(2):160–161. https://doi.org/10.1310/hpj5202-160

Garnier A, Vanherp R, Bonnabry P, Bouchoud L (2021) Use of simulation for education in hospital pharmaceutical technologies: a systematic review. Eur J Hosp Pharm. https://doi.org/10.1136/ejhpharm-2021-003034

Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P (2021) Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Mol Diversity 25(3):1315–1360. https://doi.org/10.1007/s11030-021-10217-3

Kuyt K, Park SH, Chang TP, Jung T, MacKinnon R (2021) The use of virtual reality and augmented reality to enhance cardio-pulmonary resuscitation: a scoping review. Adv Simul 6:11. https://doi.org/10.1186/s41077-021-00158-0

Patel S, Vincent AH, Abel SR, Jacobs CM, Dunlop SR, Seibert M (2011) A virtual clean room to teach USP 797 regulations for intravenous medications. Am J Pharm Educ 75(1):7

René J (2022) My doctor made me cry. It sums up everything that is wrong with health care. Today Headlines. April 10. https://goadnews.com/my-doctor-made-me-cry-it-sums-up-everything-that-is-wrong-with-health-care-today-headlines/, https://news.yahoo.com/doctor-made-cry-summed-everything-120051998.html

Rollo ME, Aguiar EJ, Williams RL, Wynne K, Kriss M, Callister R, Collins CE (2016) EHealth technologies to support nutrition and physical activity behaviors in diabetes self-management. Diab, Metab Syndr Obes: Targets Ther 9:381–390. https://doi.org/10.2147/DMSO.S95247

Roosan D, Chok J, Baskys A, Roosan MR (2022) PGxKnow: a pharmacogenomics educational HoloLens application of augmented reality and artificial intelligence. Pharmacogenomics 23(4):235–245. https://doi.org/10.2217/pgs-2021-0120

Saxena K, Diamond R, Conant RF, Mitchell TH, Gallopyn IRG, Yakimow KE (2018) Provider adoption of speech recognition and its impact on satisfaction, documentation quality, efficiency, and cost in an inpatient EHR. AMIA Summits Transl Sci Proc 2018:186–195

Schneider J, Patfield M, Croft H, Salem S, Munro I (2020) Introducing augmented reality technology to enhance learning in pharmacy education: a pilot study. Pharm: J Pharm Educ Pract 8(3):109. https://doi.org/10.3390/pharmacy8030109

Smith C, Friel CJ (2021) Development and use of augmented reality models to teach medicinal chemistry. Curr Pharm Teach Learn 13(8):1010–1017. https://doi.org/10.1016/j.cptl.2021.06.008

Suryanarayanan P, Epstein EA, Malvankar A, Lewis BL, DeGenaro L, Liang JJ, Tsou C-H, Pathak D (2021) Timely and efficient AI insights on EHR: system design. AMIA Ann Symp Proc 2020:1180–1189

Tripathi MK, Nath A, Singh TP, Ethayathulla AS, Kaur P (2021) Evolving scenario of big data and artificial intelligence (AI) in drug discovery. Mol Diversity 1–22. https://doi.org/10.1007/s11030-021-10256-w

United States: Life expectancy 1860–2020. Statista. Retrieved April 19, 2022, from https://www.statista.com/statistics/1040079/life-expectancy-united-states-all-time/

Zhao J, Xu X, Jiang H, Ding Y (2020) The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. BMC Med Educ 20:127. https://doi.org/10.1186/s12909-020-1994-z

Zhao M, Hoti K, Wang H, Raghu A, Katabi D (2021) Assessment of medication self-administration using artificial intelligence. Nat Med 27(4):727–735. https://doi.org/10.1038/s41591-021-01273-1

Chapter 14
Artificial Intelligence and Augmented Reality in Physical Activity: A Review of Systems and Devices

Jose Luis Solas-Martínez, Sara Suárez-Manzano, Manuel J. De la Torre-Cruz, and Alberto Ruiz-Ariza

Abstract In this chapter, we describe the existing AI devices aimed at PA monitoring and the most relevant AR software focusing on movement. Our review will present the latest findings, classified according to their accessibility and economic cost, and will describe the main characteristics, the basic instructions for use, and the practical applications of each of them. Within the domain of physical activity (PA), artificial intelligence (AI) has been applied primarily in the form of smart wearable devices, with the aim of improving people's living standards. Through the use of wearables (Xiaomi Smart Band, Fitbit, Apple Watch) and cameras (Alfa AI, Zenia, Peloton Guide, AndroVideo, Sportvu 2.0), AI can monitor the user's movements and physiological signals and can provide feedback about them. Augmented reality (AR) is now being employed to visualise, promote and motivate new forms of human movement. Although there are very innovative initiatives aimed at promoting the practice of PA in a virtual environment, very little is currently known about these environments. The use of AR has given rise to a myriad of proposals in all fields of motor research and development, ranging from psychomotor processes in children (FitnessMeter, AR Runner, DribbleUp) to educational/sports PA practice (Lü interactive playground, BEAM interactive floor, HADO), PA and health (PostureScreen, Complete Anatomy 2022), high-performance sport (My Jump Lab, Capture.U) and programs to promote PA in the elderly (Obie interactive projector, CyberCycle). There is software that uses both AI and AR, for example Alfa AI, Capture.U and

J. L. Solas-Martínez · S. Suárez-Manzano (✉) · M. J. De la Torre-Cruz · A. Ruiz-Ariza
Faculty of Educational Sciences, University of Jaén, Jaén, Spain
e-mail: ssuarez@ujaen.es

J. L. Solas-Martínez
e-mail: jlsm0004@red.ujaen.es

M. J. De la Torre-Cruz
e-mail: majecruz@ujaen.es

A. Ruiz-Ariza
e-mail: arariza@ujaen.es

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_14

CyberCycle. The use of these two technologies allows for the practice of PA in a new dimension, and enables personalised, real-time monitoring of a user's activity through innovative exercises.

14.1 Introduction

In recent decades, smart technology has become an essential element in the practice of physical activity (PA), as it can protect health, enable better decision making, and bring access to sport to more people. Numerous technological devices have been designed to promote health and the practice of PA, and advances in artificial intelligence (AI) and augmented reality (AR) have strong potential in terms of monitoring and motivating the practice of PA for every individual, regardless of their age and characteristics.

The aim of AI is to develop intelligent systems that mimic aspects of human behaviour, such as learning, perception, reasoning and adaptation (de Moraes et al. 2020). AI can be defined as the ability of a system to interpret and learn from external data and apply learning outcomes to achieve specific goals and solve problems through flexible adaptation (Wei et al. 2021). In the domain of PA and sport, AI enables the monitoring and recording of physiological data such as heart rate and energy expenditure, and the assessment of movement quality (Brickwood et al. 2019; Kristoffersson and Lindén 2022; Li et al. 2021). These types of information can promote the self-regulation of one's own PA, goal setting and the creation of action plans (Li et al. 2021). In addition, AI can provide feedback and create personalised training based on recorded data (Kristoffersson and Lindén 2022; Wei et al. 2021).

Also, AR can incentivise PA practice through the application of game dynamics in interactive environments. Exercise enhanced by AR is considered a new approach to exercise that can prevent non-communicable diseases within a population (Ng et al. 2019). Furthermore, the use of AR devices creates increases in PA and the motivation of users (Kosa and Uysal 2022; Nekar et al. 2022).

In view of this, the use of this technology has the potential to improve the health and well-being of the world's population. According to the WHO (2020), 28% of adults and more than 80% of adolescents have an insufficient level of PA. Physical inactivity is one of the main risk factors for chronic disease, mental health problems and increased adiposity (Cho et al. 2021; Li et al. 2021; WHO 2020). In order to combat this problem, new methods are being sought to encourage the practice of PA, such as the creation of programs or interventions aimed at its promotion (Creaser et al. 2021; Winand et al. 2020) and more concrete, accessible and attractive guidelines for the general population (Smith and Wightman 2021).

One of the key factors that can explain a lack of sustained PA over time is a lack of motivation (Winand et al. 2020). Other obstacles that have been found to prevent people from engaging in PA are social factors such as the cost of activities, the availability of transport, a lack of family support, and limited time availability (West et al. 2021). For this reason, there is a need to explore new technology-based intervention methods that are effective, accessible, and offer greater availability (Ni et al. 2019; Winand et al. 2020). Recent studies have claimed that technological interventions could lead to higher levels of PA and active participation in sport (Cho et al. 2021; Ni et al. 2019).

In order to bring this technology closer to the general population and to encourage its use in PA practice, we review and classify the AI devices intended for PA monitoring and the most relevant AR software focusing on movement and motor development. The latest findings are classified based on the type of device, its accessibility and the economic costs involved. We also describe the main features and basic instructions for use, as well as some of the practical applications of each device.

14.2 Artificial Intelligence Devices for Physical Activity Monitoring

AI may become very important in the field of PA, due to the relative difficulty of accurately analysing activity without expensive equipment. In conjunction with a lack of sports training of most people, this could lead to poor decision making that negatively affects their health (Wei et al. 2021). However, this can be avoided through the use of AI, which is capable of monitoring and supervising all training with the aim of giving information to the user and can even analyse all of the collected data to provide recommendations to the user on how to progress in their training (Ferguson et al. 2021; Wei et al. 2021). Furthermore, with these devices, the cost of personal trainers or gym memberships and the time spent travelling to fitness centres could be reduced (D'Amore et al. 2021; Oh et al. 2021).

Advances in this field allow sportspeople to improve their performance and avoid injury, thanks to the information provided by AI on breathing, heart rate and effort level. The most commonly used devices are wearables such as the *Xiaomi Smart Band* or *Apple Watch*. These devices synchronously record real-time physiological data such as daily steps, heart rate and energy expenditure (Brickwood et al. 2019). In addition to these, AI-enabled camera systems (smartphone or other specific devices) can be employed that are capable of collecting a wide range of biomechanical data from the user. In the following sections, we review the functionality of several types of devices and some of the most commonly used mobile applications offered for PA monitoring.

14.2.1 Wearables

Wearables are small portable devices that can take the form of accessories or clothing. They are designed to use modern information networks and various sensor technologies to collect, transmit and analyse the physiological and health parameters of the

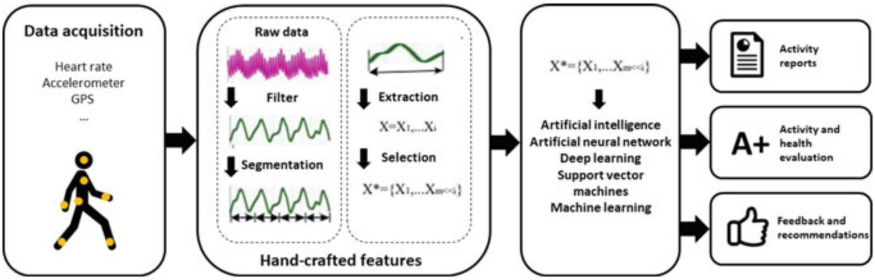


Fig. 14.1 Scheme showing how wearables obtain and process information

human body, in order to monitor the overall physical state of an individual (Nahavandi et al. 2022; Tang et al. 2020; Wei et al. 2021). Some of these parameters include energy expenditure, distance travelled, heart rate and sleep.

These devices can monitor the entire physical training process and can provide users with actionable advice through a big data analysis system. For example, wearables can accurately measure user-specific data during exercise, such as heart rate, and can issue warnings when a safe heart rate is exceeded. A diagram illustrating how the information collected by wearables is processed before being presented to the user is shown in Fig. 14.1.

Wearables are affordable, visually appealing, easy to use, and encourage an active lifestyle (Ferguson et al. 2021; Li et al. 2021; Tang et al. 2020). There is evidence that wearable devices are an effective way to increase steps and moderate-to-vigorous PA and can reduce sedentary behaviour in both young people (Creaser et al. 2021) and adults (Brickwood et al. 2019). Although there is a great variety of wearables, such as smart glasses or smart clothes, the most common are smart watches or smart bands, due to their accessibility and ease of use. In the following, we describe some current wearables, along with their main features and functions.

***Xiaomi Smart Band**:* This model is one of the most accessible on the market. It has a wide variety of functionalities and sensors that allow it to monitor a multitude of biometric markers including heart rate, sleep quality, stress monitoring and breathing. From these data, the user can see a summary of his or her activity; however, for more detailed information on all metrics, the device must be linked to the *Zepp Life* app. It records all biometrics and provides feedback on sleep and workouts, and also offers video tutorials that aim to motivate the user and help to create a more active and healthy daily routine (Fig. 14.2).

This smart band model is designed for any user and aims to monitor some important variables for health.

Fitbit Charge 5: This has more functionalities and more sophisticated sensors than the *Xiaomi Smart Band*. *Fitbit* is able to perform electrocardiograms and to monitor electrodermal activity and oxygen saturation. In addition, this model has an integrated GPS, which is ideal for runners. This device also requires synchronisation with the

Fig. 14.2 Xiaomi Smart Band 6 performing a blood oxygen saturation (SpO_2) analysis (extracted from https://www.youtube.com/watch?v=DiNZT-wL89E&t)

Fitbit app to allow the user to see all of the relevant statistics and to access more functionalities; challenges, health and training programs are also provided (Fig. 14.3).

This smart band is more complete than the *Xiaomi Smart Band*, and its price is therefore higher. It is designed for more active users who want to monitor their health status with great accuracy.

Apple Watch Series 7: This smart watch has a large number of sophisticated biosensors, and a wide range of functionalities and applications that can be used without the need to link to a smartphone. It can be used to browse all the data collected, and various kinds of applications with a focus on the practice of PA and health can be installed (Fig. 14.4).

The *Apple Watch* is not a device with the sole purpose of monitoring activity and health, and also has many productivity and entertainment applications that can be accessed at any time from the wrist.

Fig. 14.3 Fitbit Charge 5 performing an electrocardiogram (extracted from https://www.youtube.com/watch?v=W0KfJCUEcTA&t)

Fig. 14.4 *Smart Watch Series 7* displaying daily activity data (extracted from https://www.youtube.com/watch?v=mpG1VWxOOec&t)

To choose the wearable that best suits the user's needs, several aspects must be taken into account, such as the type of sensors included, the autonomy of the device, and its price. Table 14.1 shows the specifications and price of each of the wearables described above.

14.2.2 Artificial Intelligence Camera Systems

AI camera systems consist of cameras and computers that are capable of capturing and automatically identifying people and analysing their movements by comparing them with the information available to the AI, in order to provide feedback on exercise performance (Wei et al. 2021). Although the information in a video can help users to analyse their performance, users can correct and assess all their movements more accurately with the help of AI analysis. A diagram of how AI camera systems obtain and process data before presenting information to the user is shown in Fig. 14.5.

The AI system detects each individual and their body structure. Following this, the system collects information about the execution of each exercise and compares it with the references stored in its database. It then calculates the level of coincidence and locates any errors in the execution. The system is able to estimate the amount of PA performed by individuals and to calculate the calories consumed (Cruz et al. 2021). Moreover, users can browse all the statistics related to their training and observe which aspects can be improved.

AI camera systems have two major advantages: the first is the comfort they offer the user when performing movements, since these systems are not invasive or uncomfortable, while the second is the amount of information that can be displayed in image or video format, making it easy to understand by any user (Mabrouk and Zagrouba 2018). However, this type of system is very expensive, and is not very feasible to

Table 14.1 Features and approximate price of some of the smart bands and smart watches on the market

Device		Features	Price
Xiaomi	*Smart Band 4*	• Designed for indoor and outdoor exercise • Up to 20 days of autonomy • Waterproof up to 50 m • Six training modes • Heart rate monitoring • Sleep monitoring (light and deep) • *Zepp Life* app can be used to check health and daily PA data	Under 20$
	Smart Band 5	• Up to 14 days of autonomy • Waterproof up to 50 m • 11 training modes • Intelligent 24-h heart rate monitoring with notification of any abnormalities • Sleep monitoring (light, deep and REM) • Three health modes to help users improve their health • Intelligent personal activity monitoring • Stress monitoring • Breathing exercises • *Zepp Life* app can be used to browse daily health and PA data, and special exercises are offered to meet individual needs	From 20$
	Smart Band 6	Smart Band 5 features plus: • 30 training modes • Blood oxygen saturation monitoring • Sleep monitoring of naps during the day and breathing quality during sleep	From 30$
	Smart Band 7	Smart Band 6 features plus: • More than 120 sport modes • Training load analysis	From 40$
Fitbit	*Fitbit Charge 5*	• 20 exercise modes • Waterproof up to 50 m • Up to seven days of battery life • Electrocardiogram application • Stress management level • Oxygen saturation monitoring (SpO_2) • Detailed heart rate monitoring and analysis with display of health and wellness statistics • Monitoring of HR, recovery level and time in Active Zone • Respiratory rate • Skin temperature monitoring • Electrodermal activity scanner • Sleep score and sleep phases • Integrated GPS	From 100$

(continued)

Table 14.1 (continued)

Device		Features	Price
Apple	*Apple Watch Series 7*	• Apps provided by the smartwatch itself – A large number of applications can be installed or linked – Training design: strength, aerobic, anaerobic, HIIT, etc. – Training control for various activities such as running, cycling, swimming, etc. • Intelligent monitoring of heart rate, breathing and sleep quality • Wi-Fi connection independent of smartphone connection • 18 h of battery autonomy and 80% battery charge in only 45 min • Monitors blood oxygen level • Checks the user's heart rate and can perform electrocardiograms • Sleep monitoring • Monitors respiratory rate even during the night • Integrated GPS	From $300

PA: physical activity; *HIIT*: high-intensity interval training

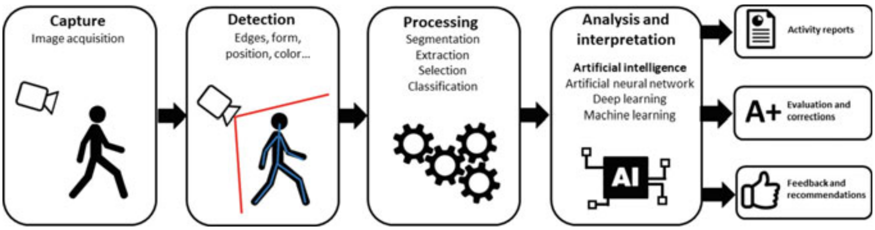


Fig. 14.5 Scheme showing the data acquisition and processing of AI camera systems, from image capture until the information is displayed to the user

install in a sports venue. Nevertheless, in places such as gyms or fitness centres, access to such devices may be valuable for performance improvement and injury prevention. In addition, it should be noted that the accuracy of these systems can be reduced by changes in lighting, frequent alterations in the scene background, the type of clothing worn by the subject and other disturbances (Mabrouk and Zagrouba 2018; Wang et al. 2019). Overall, however, AI camera systems are good tools for encouraging and improving PA practice, both for beginners (cheaper devices) and for advanced athletes (more expensive devices).

In the field of PA, this technology is mainly used to control the execution of exercises and to monitor a wide range of statistics for the entire training process. In this chapter, we review four of the most widely used AI camera systems in use today, from the most accessible to the least.

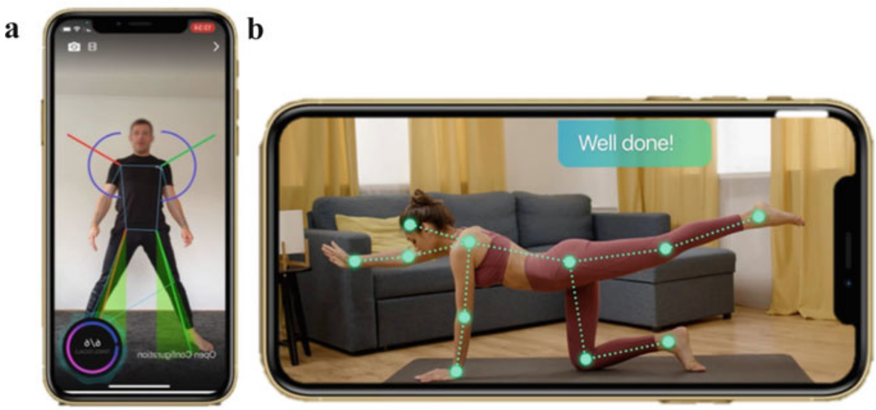


Fig. 14.6 **a** Image captured by *Alpha AI* during the performance of jumping jacks (taken from https://www.youtube.com/watch?v=SvuZ9MacW2g); **b** image captured by the *Zenia* app to evaluate a yoga pose (taken from https://www.youtube.com/watch?v=ZH4BbR8Zi-A).

(1) ***Alfa AI*** and ***Zenia*** are smartphone fitness apps that are capable of "replacing" personal trainers. These apps offer training sessions that monitor and provide real-time feedback on each exercise for optimal training. *Alfa IA* is dedicated to fitness exercises such as planks, push-ups and climbers, while *Zenia* focuses on yoga (Fig. 14.6). The biggest advantage of these apps is that they can be used anytime, anywhere. In addition, there are a wide variety of applications focused on PA tracking that are similar to these, with different functionalities and features, some of which are free and others are paid.

These applications are intended for personal use, and anyone can purchase them to start performing PA on a regular basis. For a more successful experience, personal motivation and some experience in sports practice are recommended. Both require monthly subscriptions starting at $5.

(2) ***Peloton Guide*** consists of a camera system connected to a television, which shows how to perform training, monitors the execution of movements, and provides feedback in real time (Fig. 14.7). This system provides recommendations for programs and workouts that can improve a training routine based on the data collected during each session. Another well-known example of a device that is similar to this technology is *Kinect* for Xbox.

This device is aimed at the general public, for personal use. It is not as accessible as a wearable, but it is capable of creating and presenting instructing workout routines while monitoring the execution of each exercise, in the same way as a personal trainer. Currently, it requires an outlay of approximately $300.

(3) ***AndroVideo*'s** AI camera systems detect the position of the user's body, provide real-time posture correction, automatically record the user's condition, display post-workout recommendations, and create a history log (Fig. 14.8). This is device is aimed at users of gyms or fitness centres. It is similar to *Peloton*

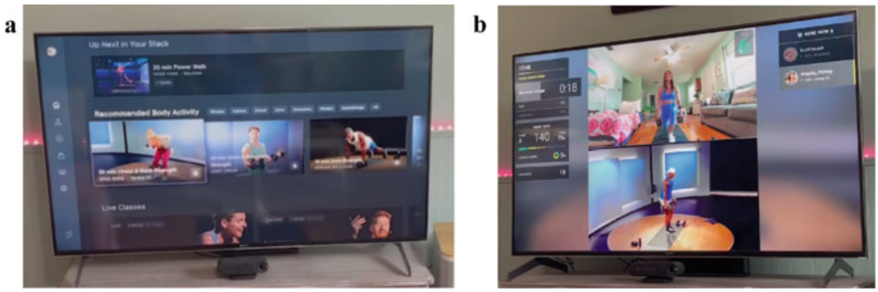


Fig. 14.7 **a** Main menu of *Peloton Guide* (extracted from https://www.youtube.com/watch?v=oZMXpoikDXs&t=757s); **b** screen capture of *Peloton Guide* during lunges (https://www.youtube.com/watch?v=sHGzPtVHk8Y)

Guide but acts as an artificial assistant rather than a personal trainer. It focuses on posture correction, exercise guidance, and condition monitoring. Using this device, trainers and users can access their exercise history and their statistics, in order to adapt their training plan according to their needs and goals.

(4) ***SportVU 2.0*** is a multi-camera system that collects high quality images of complex game situations, in order to provide accurate tactical (player positions) and physical (physical condition of each player) data. This system is used in elite sports such as the NBA and professional soccer leagues. Its function is to follow the ball and all of the players, and to provide statistics on their positions on the court in real time (Fig. 14.9). All of this information is intended to be used by teams to improve their performance. Using these systems, game broadcasts can show data about the game, such as the percentage of possession, the percentage of successful three-pointers, the maximum speed reached or the distance travelled by each player. However, this is a device aimed exclusively at elite teams with high budgets; for example, for use in the NBA, *SportVU 2.0* charges about $100,000 per team.

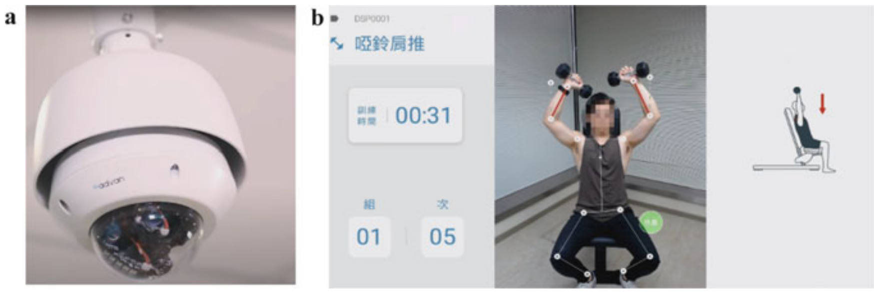


Fig. 14.8 **a** *AndroVideo* AI camera (extracted from https://www.youtube.com/watch?v=PtSGp6JGP_E); **b** *AndroVideo* screen, indicating in red that the user is performing a military dumbbell press incorrectly (extracted from https://www.youtube.com/watch?v=N19usqnNdHM)

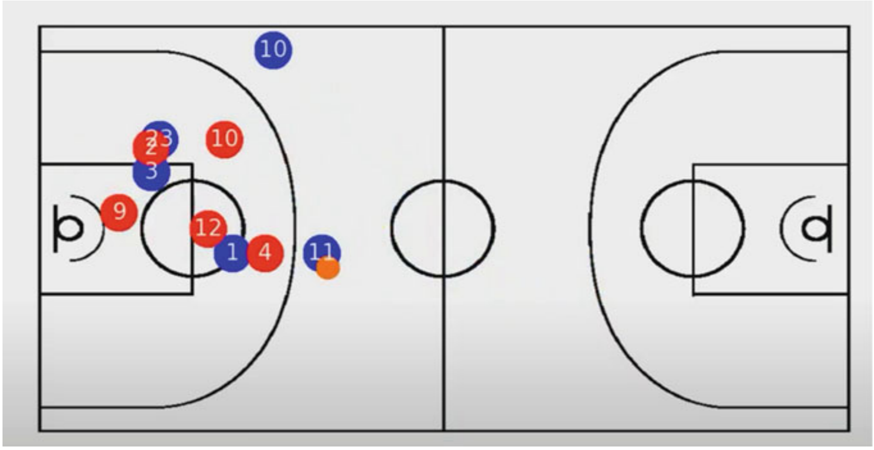


Fig. 14.9 Example of a situation in a real basketball game mapped with *SportVU 2.0* AI (extracted from https://www.youtube.com/watch?v=Yn_7bb_ruEw)

14.3 Augmented Reality Software for Movement and Motor Development

AR is a technology that involves generating new images, sounds or text from digital information in a real physical environment, thus simulating an environment in which the artificial and the real are mixed (Gil et al. 2021). With AR, the surrounding world becomes interactive and digitally manipulable. The use of AR opens up a wide range of possibilities in fields of research related to movement and motor development, from psychomotor processes in children to the promotion of PA in the elderly. Its exponential growth in recent years has also allowed increasing numbers of people to access this technology through smartphones, computers and other devices (Nekar et al. 2022).

Through AR, children can develop their abilities to perceive, recognise and process various objects and situations; there is also the possibility of adding, removing or modifying virtual aspects with the aim of personalising the experience according to their needs and abilities (Soltani and Morice 2020). In addition, several studies have revealed that AR games can increase daily PA levels and time spent outdoors (Bonus et al. 2018; Kosa and Uysal 2022; Ruiz-Ariza et al. 2018) and can stimulate players' development and cognition (Ruiz-Ariza et al. 2017; Soltani and Morice 2020).

AR systems feed augmented information to players' sensory modalities to help them understand the real-world environment (Soltani and Morice 2020). Mobile AR games have the potential to make environmental exploration enjoyable (Winand et al. 2020) and to create positive life experiences (Bonus et al. 2018). There are currently a number of AR games on the market, the most famous example of which is *Pokémon GO* (Ruiz-Ariza et al. 2018). In this game, AR is used to project images of Pokémon onto images of the real environment, using a cell phone camera. In both youths and adults, increases in moderate PA and numbers of steps were detected up to seven months from its first use (Baranowski and Lyons 2020).

Existing evidence indicates that PA practice and motivation can be improved through AR devices (Gil et al. 2021). Training with AR can not only replace traditional training but can be even more effective in developing certain skills or abilities such as balance, strength or muscular endurance (Nekar et al. 2022). At the same time, the experience of AR games contributes to well-being, as being voluntarily present in a virtual environment can elevate mood, regulate emotions and provide a healthy escape (Kosa and Uysal 2022).

In the domain of high-performance elite sport, AR is used by athletes, coaches and researchers. This technology can mine biomechanical data to project simulations onto the users of the movements they are performing. All of the information provided is used for movement correction and to improve the athlete's technique (Bolam et al. 2021). In addition, it can also be used as a validated assessment tool for various physical tests, such as jumping or running speeds (Bishop et al. 2022).

For the elderly, AR systems serve as a tool that can encourage rehabilitation through fun and entertainment (Jeon and Kim 2020). In studies of AR, seniors were interested and motivated by exercise, and exercise control was found to be easy for them (Jeon and Kim 2020). Table 14.2 briefly describes the main features of the AR applications and devices reviewed in this chapter, and they are described in more detail below.

14.3.1 Psychomotor Processes in Children

FitnessMeter: This is an advanced timing and measurement tool that can be used to perform various physical fitness tests. It is aimed at both coaches and physical education teachers, to allow them to evaluate their pupils and for personal use. It is capable of measuring speed, agility, jump height and general fitness through various tests: sprinting over 10 or 20 m, vertical jumping, the Course Navette test, and repetitions of push-ups, sit-ups or other exercises (Fig. 14.10).

AR Runner: This is an AR game which uses the camera to create a series of virtual control points over the terrain, which must be traversed as quickly as possible. All of the information is displayed on the phone screen (Fig. 14.11), and it is possible for a user to compete with other players from all over the world.

Table 14.2 Description, main target audience and approximate economic cost of apps and devices that employ AR

Field of work	Device/app	Description	Main target audience	Economic cost
Psychomotor skills in children	*FitnessMeter*	Fitness testing and physical performance evaluation	Any user	$2
	AR Runner	Competitive AR game in which the user races through checkpoints	Any user	Free
	DribbleUp	Intelligent fitness equipment connected to an app, with the aim of developing sports skills through different challenges or activities	Teenagers, young adults or amateur athletes	$49.99 + $16.99/month subscription
Educational/sports practice of PA	*Lü* interactive projector	Interactive projector that transforms any indoor space into an active and immersive educational environment in which children can participate physically, intellectually and socio-emotionally in different activities	Schools, institutes or youth centres	$20,000
	BEAM interactive floor	Play system with an interactive projector focused on the floor, which provides different active and educational games for children to learn and have fun	Schools, institutes or youth centres	$10,000

(continued)

Table 14.2 (continued)

Field of work	Device/app	Description	Main target audience	Economic cost
	HADO	New AR e-sport, which is played alone or in a team; equipped with an AR helmet and a connected armband to detect movements, the player can see virtual elements added to the real world and interact with them	Educational, sports or leisure centres	Rental from an official HADO site: 100$/hour or 10$/person for three games Installation cost unavailable
PA and health	*PostureScreen*	Assessment software that analyses and evaluates posture and body composition	Health professionals assessing posture and fitness	$50
	Complete Anatomy 2022	Complete anatomy platform with unique learning and collaboration tools; AR function available for mobile app	Students or health professionals	$75 for students $120 for professionals
High-performance sports	*Capture.U*	System that collects data through *Blue Trident* inertial sensors; displays real-time data overlaid on the video, allowing the user to make informed decisions	Amateur and elite sports professionals	From $6600

(continued)

Table 14.2 (continued)

Field of work	Device/app	Description	Main target audience	Economic cost
	My Jump Lab	Allows the user to carry out more than 30 tests, including jumps, lifts, sprints, changes of direction or even motion capture with great accuracy	Elite athletes and sport professionals	Free + $3.99/month subscription
Promotion of PA in the elderly	*Obie for Seniors* virtual interactive projector	Interactive game console that projects custom-made images onto any surface, encouraging active play through touch, movement and hand–eye coordination via the displayed images	Elderly people	From $3300
	CyberCycle	Type of exergame that combines a traditional stationary bike with virtual reality rides, competitive avatars and video game features	Elderly people	From $6000

PA: physical activity; *AR*: augmented reality

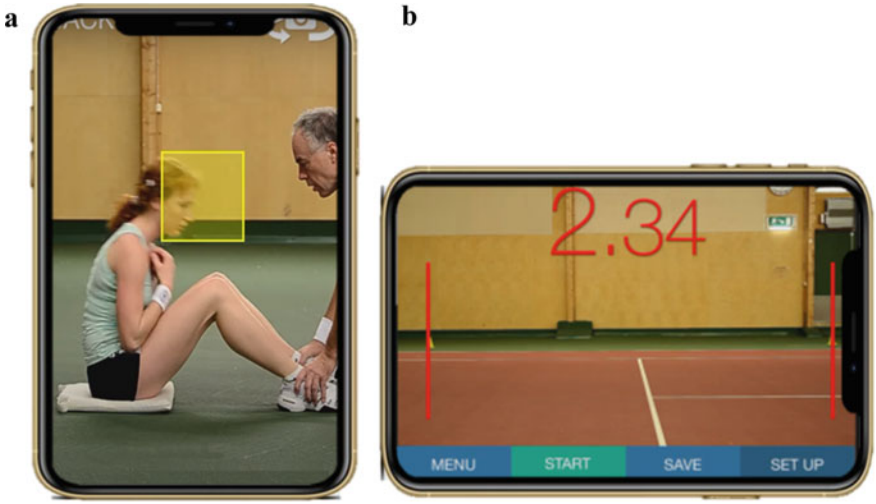


Fig. 14.10 **a** Abs test with *FitnessMeter*; **b** speed test with *FitnessMeter* (images extracted from https://www.youtube.com/watch?v=KuK1kTb7hp8)

Fig. 14.11 **a** Main menu of *AR Runner*; **b** screen during a game of “30 s rush” in *AR Runner* (images extracted from https://www.youtube.com/watch?v=5QouVbIjXgM&t)

Fig. 14.12 **a** Soccer technique training with *DribbleUp* (extracted from https://www.youtube.com/watch?v=1wYhFLVhp3M&t); **b** basketball dribble training with *DribbleUp* (extracted from https://www.youtube.com/watch?v=B3VeT7H62kQ)

***DribbleUp*:** Smart equipment such as soccer balls, basketballs, medicine balls and boxing gloves can be linked to a smartphone via the *DribbleUp—Sports & Fitness* app. The app collects metrics such as crossing speed, dribbling motion, resistance, repetitions and other variables, via the cell phone camera and equipment sensors (Fig. 14.12). The objective is to improve the user's technique in different sports skills through exercises that are presented in the form of challenges, in the same way as a video game.

14.3.2 Educational/Sports Physical Activity

***Lü* interactive playground:** *Lü* is an intelligent spatial environment that transforms a gymnasium or other indoor space into an interactive play area. It creates a virtual learning zone where young people can move and perform PA, in an attractive and innovative way.

This system consists of one or two video projectors, a computer, a 3D video camera, six static lights, two computerised lights and two loudspeakers. The complete system is able to project interactive games onto the walls, thus creating an interactive space in which young people can perform different innovative physical activities (Fig. 14.13). *Lü* offers various activities that encourage teamwork, competition and self-improvement through point systems. The different games that are available include subject matter such as mathematics, science, geography and foreign languages.

***BEAM* interactive floor:** This is an interactive projection system that is focused on the floor, and allows several people to interact with it at the same time. It offers more than 200 educational games through which young people can learn via motion. *BEAM* can not only develop gross and fine motor skills, but also works on different aspects of cognition (memory, attention or perception). This system provides children and adolescents with the opportunity to explore, discover and interact through movement and fun (Fig. 14.14).

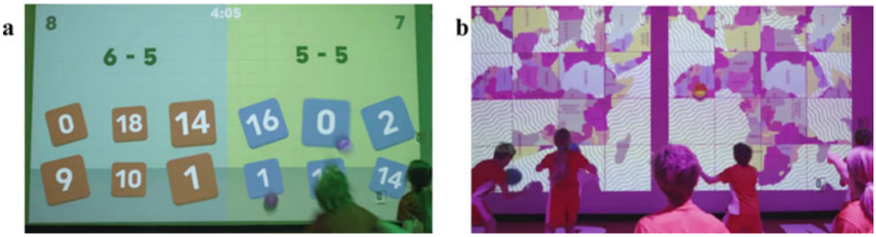


Fig. 14.13 Educational and active competitive team games projected on a wall with *Lü*: **a** active mental arithmetic game; **b** active game with geography content (images extracted from https://www.youtube.com/watch?v=n-gbscvwkGA)

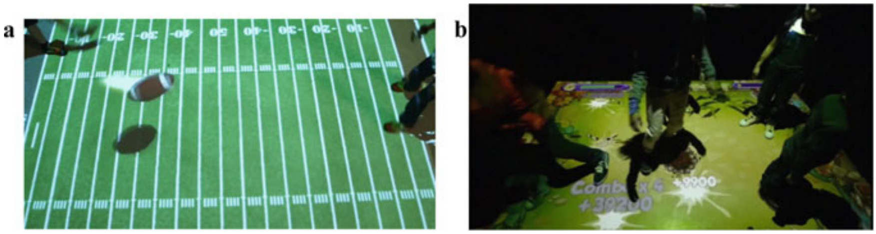


Fig. 14.14 Active games projected on the ground using *BEAM*: **a** competitive game of American soccer (extracted from https://www.youtube.com/watch?v=Z3rbKO1pv0U); **b** cooperative game of catch the mole (extracted from https://www.youtube.com/watch?v=SbCHRL-E4qY)

HADO: This is a new AR e-sport that can be played individually or in teams. Players are equipped with two smartphones: one is placed in a helmet to function as AR glasses, and the other is positioned in an armband to act as a motion sensor (Fig. 14.15). Players can see and interact with virtual elements added to the real world.

The main game offered by the *HADO* system is a competitive e-sport between two teams of three players. Each game lasts 80 s, and the team with the most points wins. Each team member must throw energy balls at the opposing team's players; if the opponent's body is hit, he or she loses one life cell. Each player has four life cells, and if all of these are lost, the other team gains one point. In addition to throwing energy balls, each player can create an energy shield for protection, and the balls can also be dodged. A bar is shown for each player, which indicates the level of energy available and limits the use of energy balls and shields (Fig. 14.16).

HADO is a very successful e-sport in Japan and is becoming increasingly widespread throughout the world. It is a new way of practicing sports that can be very attractive to the younger generation. It is an innovative device that illustrates the wide range of possibilities that AR can offer to the practice of PA, in terms of both promoting its practice and developing different motor skills.

Fig. 14.15 HADO equipment. Helmet and armband with attached smartphones (extracted from https://www.youtube.com/watch?v=GSXUtScLzEo)

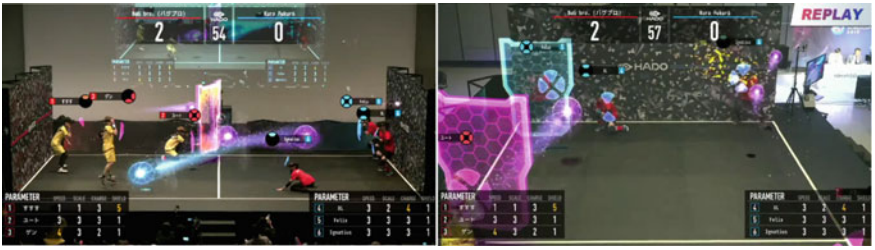


Fig. 14.16 *HADO* quarterfinal match at the 2019 World Cup (extracted from https://www.youtube.com/watch?v=GSXUtScLzEo)

14.3.3 Physical Activity and Health

***PostureScreen*:** This is a smartphone app that is available for both Android and IOS. It is an objective assessment tool for posture and movement that is capable of presenting and automatically emailing the results to the patient. It also offers customised, safe and functional postural exercises through another application called *WebExercises* to which it is linked (Fig. 14.17).

The use of this app for postural assessment is supported by scientific studies, such as those by Boland et al. (2016) and Szucs and Brown (2018), which report that the app demonstrates strong reliability and validity and can be employed in both clinical and research settings.

***Complete Anatomy 2022*:** This is a 3D anatomy mobile application platform with unique collaboration and learning tools. The user can interact with the model by selecting an element (muscle, organ, blood vessel etc.) and exploring and labelling

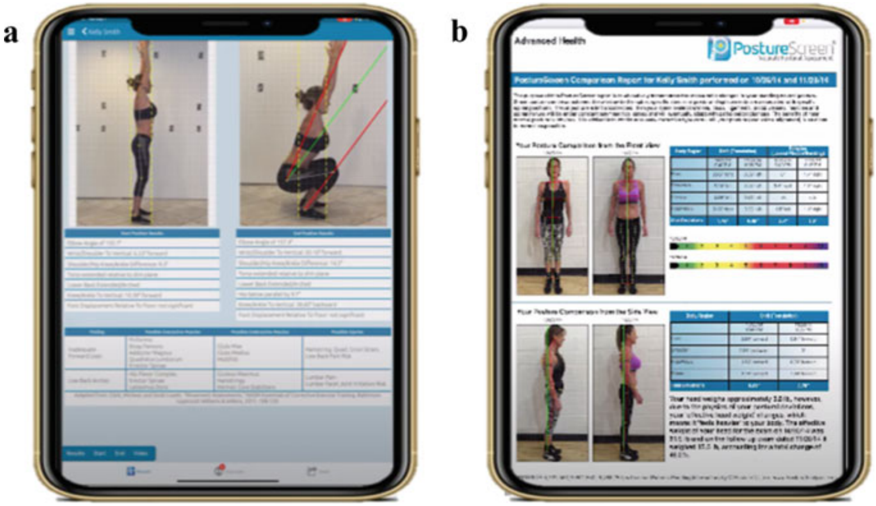


Fig. 14.17 **a** Report from *PostureScreen* on the execution of a squat; **b** report comparing the posture of the same person at different times (images extracted from https://www.youtube.com/watch?v=T_TLC91Ay_o&t)

the structures in real time. The program simulates the condition and details of the human body, allowing a user to explore human anatomy in depth using the cell phone or tablet as a "viewer" (Fig. 14.18).

The app has a set of tools that can cut out the layers that make up the body, allowing the user to explore the relationships between structures (muscular, skeletal,

Fig. 14.18 Images from the *Complete Anatomy 2022* mobile app, showing the composition of the human body at the muscular and vascular level (images extracted from https://www.youtube.com/watch?v=CjD-nFk7qtM)

vascular), simulate injuries and pathologies, animate pain points in the model, and add labels directly to a selected structure.

14.3.4 High-Performance Sports

Capture.U: The *Vicon* application collects and analyses biomechanical information from the human body in real time, and displays a simulation overlaid on the real image. This allows the user to make informed decisions in different situations, such as on the playing field, on the track, in the pool, on the court or in the lab. For full data collection, the user is required to wear *Blue Trident* sensors.

This application has a visualisation mode based on AR, in which 2D, 3D or both types of visualisations are shown at the same time (Fig. 14.19). There is also the option to select a particular joint of interest to view its kinematic data, positions and opening angles.

Some studies, such as that of Bolam et al. (2021), report that this application can be used both in rehabilitation for movement correction and in high-performance sports to improve execution.

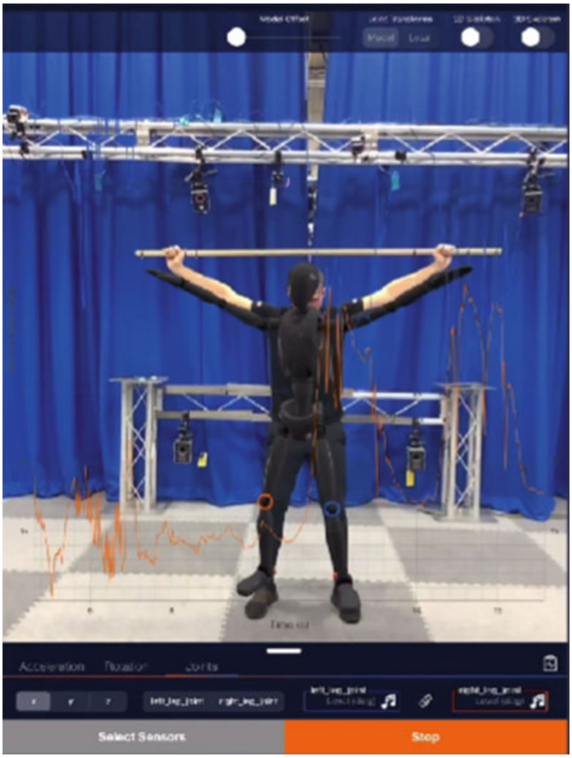


Fig. 14.19 3D model visualisation of squat movements in *Capture.U*, with a graph of the frequency and amplitude of the movements superimposed on a real-time image (extracted from https://www.youtube.com/watch?v=kRuio13OqQk)

Fig. 14.20 *My Jump Lab*: **a** main menu (extracted from https://www.youtube.com/watch?v=mObR35P_gm8&t); **b** vertical jump test (taken from https://www.youtube.com/watch?v=mObR35P_gm8&t); **c** speed test for each dead weight repetition (taken from https://www.youtube.com/watch?v=dKUEHvqKk5E)

***My Jump Lab*:** This is an IOS application that performs the function of a portable and validated performance evaluation laboratory. No additional accessories are required for its use. Simply by using the smartphone's camera and sensors, the app is able to accurately record data on different physical skills (Fig. 14.20). With *My Jump Lab*, more than 30 tests can be performed, including jumping (counter movement jump, squat jump, repeated jumps, etc.), running and sprinting (contact/flight times, leg asymmetry, pronation/supination, etc.) and range of motion (dorsiflexion, neck extension/flexion, hip flexion and internal rotation, and shoulder flexion and internal rotation). This app is used by coaches, elite athletes, physiotherapists and researchers around the world. Its accuracy was validated by Bishop et al. (2022), who indicated that the app was a good measuring instrument for all the physical tests offered.

14.3.5 Promotion of PA in the Elderly

***Obie* interactive surface:** This is a projector that can turn any floor, wall or table into an interactive surface. It uses sensors to detect movements such as bumps, taps, or shakes, and has more than 200 games to stimulate PA or cognitive skills (Fig. 14.21). The animated games are customised to different levels, and may include memory tests, musical contests, or speed coordination challenges.

***CyberCycle*:** This is an interactive exercise bike that includes a virtual environment in which games, group rides, and competitions can be held to engage users in PA. A 19-inch HD monitor is used to immerse users in a virtual race, in which they

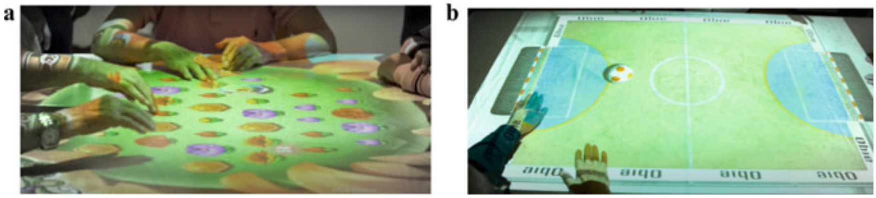


Fig. 14.21 *Obie* interactive surface: **a** collaborative memory game; **b** soccer competition game (images extracted from https://www.youtube.com/watch?v=Ml2pX8fiZJE)

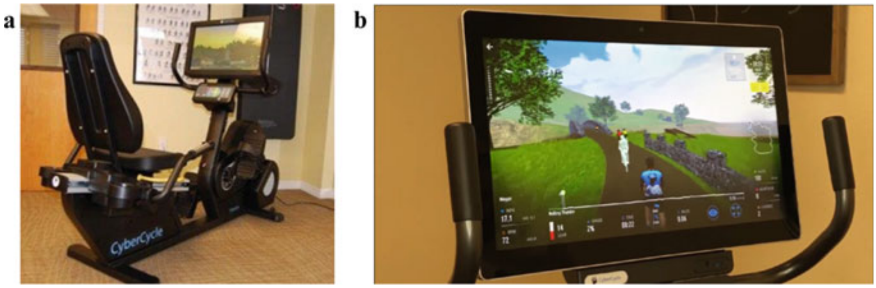


Fig. 14.22 **a** *CyberCycle* (extracted from https://www.prioritymarketing.com/gulf-coast-villages-new-cybercycle-takes-senior-fitness-to-a-virtual-level/); **b** screen capture during a virtual race (extracted from https://www.youtube.com/watch?v=acxXm10xsZ4&t)

ride through a 3D scenario (Fig. 14.22). This system is particularly suitable for the elderly, as it provides stimulation and keeps both body and mind fit.

14.4 Conclusions

AI and AR devices can encourage and enhance PA, as they increase motivation levels, monitor activity and optimise training. The use of either these technologies has great advantages for users, but the benefits are further enhanced by employing both simultaneously; this takes the practice of PA to a new level, as it allows for personalised real-time monitoring of activity via the big data processing enabled by AI and encourages PA with innovative activities offered through AR. For example, wearables can be used to monitor vital signs while exercising with AR systems (*AR Runner*, *DribbleUp* or *Lü* interactive playground), and the intensity with which the activities were performed can be checked at a later stage. Some of the devices and applications described above (*Alfa AI*, *Zenia*, *Capture.U* or *DribbleUp*) already use both technologies, where AI analyses the movements and presents feedback via AR, thus providing very complete information allowing the user to understand and correctly follow the instructions.

Emerging technologies such as AI and AR provide new opportunities for monitoring and promoting PA on a larger scale, for both personal and group use. Promoting the use of wearables and AR applications for smartphones or other devices could not only generate health and performance benefits at the individual level but could also improve the collective health of the population (Gal et al. 2018). Likewise, interventions aiming to promote PA can incorporate devices based on AI and AR, with the objective of expanding the possibilities and generating interest for a greater number of people.

Finally, as observed in this chapter, AI and AR are used in many fields of PA, from the promotion of PA at all ages to high-performance sports. In view of this, and due to the increasing accessibility of equipment with a multitude of highly reliable functionalities, the use of this technology has the potential to completely change the nature of PA and to create a more physically active society.

References

Baranowski T, Lyons EJ (2020) Scoping review of Pokémon Go: comprehensive assessment of augmented reality for physical activity change. Game Health J 9(2):71–84. https://doi.org/10.1089/g4h.2019.0034

Bishop C, Jarvis P, Turner A, Balsalobre-Fernandez C (2022) Validity and reliability of strategy metrics to assess countermovement jump performance using the newly developed my jump lab smartphone application. J Hum Kinet 83(July):185–195. https://doi.org/10.2478/hukin-2022-0098

Bolam SM, Batinica B, Yeung TC, Weaver S, Cantamessa A, Vanderboor TC, Yeung S, Munro JT, Fernandez JW, Besier TF, Monk AP (2021) Remote patient monitoring with wearable sensors following knee arthroplasty. Sensors 21(15):1–12. https://doi.org/10.3390/s21155143

Boland DM, Neufeld EV, Ruddell J, Dolezal BA, Cooper CB (2016) Inter- and intra-rater agreement of static posture analysis using a mobile application. J Phys Ther Sci 28(12):3398–3402. https://doi.org/10.1589/jpts.28.3398

Bonus JA, Peebles A, Mares ML, Sarmiento IG (2018) Look on the bright side (of media effects): Pokémon Go as a catalyst for positive life experiences. Media Psychol 21(2):263–287. https://doi.org/10.1080/15213269.2017.1305280

Brickwood KJ, Watson G, O'brien J, Williams AD (2019) Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. JMIR mHealth uHealth 7(4). https://doi.org/10.2196/11819

Cho I, Kaplanidou K, Sato S (2021) Gamified wearable fitness tracker for physical activity: a comprehensive literature review. Sustainability (switzerland) 13(13):1–15. https://doi.org/10.3390/su13137017

Creaser AV, Clemes SA, Costa S, Hall J, Ridgers ND, Barber SE, Bingham DD (2021) The acceptability, feasibility and effectiveness of wearable activity trackers for increasing physical activity in children and adolescents: a systematic review. Int J Environ Res Public Health 18(12). https://doi.org/10.3390/ijerph18126211

Cruz BS, Aguía K, Perdomo OJ (2021) Monitoring and evaluation of people in indoors and outdoors using deep learning. 78. https://doi.org/10.1117/12.2606334

D'Amore C, Reid JC, Chan M, Fan S, Huang A, Louie J, Tran A, Chauvin S, Beauchamp MK (2021) Smart technology vs. face-to-face physical activity interventions in older adults: a systematic review protocol. JBI Evid Synth 19(10):2801–2812. https://doi.org/10.11124/JBIES-21-00072 de Moraes, MHB, Ferreira DD, Ferreira AC, da Silva GR, Caetano AS, Braz, VN (2020) Use of artificial intelligence in precision nutrition and fitness. In: Debmalya B (ed) Artificial intelligence in precision health. Elsevier. https://doi.org/10.1016/b978-0-12-817133-2.00020-3

Ferguson T, Olds T, Curtis R, Blake H, Crozier AJ, Dankiw K, Dumuid D, Kasai D, Connor EO, Virgara R, Maher C (2021) Review effectiveness of wearable activity trackers to increase physical activity and improve health: a systematic review of systematic reviews and meta-analyses. Lancet Dig Health 4(8):e615–e626. https://doi.org/10.1016/S2589-7500(22)00111-X

Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM (2018) The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 4(1):1–15. https://doi.org/10.1186/s40798-018-0157-9

Gil MJV, Gonzalez-Medina G, Lucena-Anton D, Perez-Cabezas V, Ruiz-Molinero MC, Martín-Valero R (2021) Augmented reality in physical therapy: systematic review and meta-analysis. JMIR Serious Games 9(4):1–20. https://doi.org/10.2196/30985

Jeon S, Kim J (2020) Effects of augmented-reality-based exercise on muscle parameters, physical performance, and exercise self-efficacy for older adults. Int J Environ Res Public Health 17(9). https://doi.org/10.3390/ijerph17093260

Kosa M, Uysal A (2022) Effects of presence and physical activity on player well-being in augmented reality games: a diary study. Int J Hum-Comput Interact 38(1):93–101. https://doi.org/10.1080/10447318.2021.1925437

Kristoffersson A, Lindén M (2022) A systematic review of wearable sensors for monitoring physical activity. Sensors 22(2). https://doi.org/10.3390/s22020573

Li C, Chen X, Bi X (2021) Wearable activity trackers for promoting physical activity: a systematic meta-analytic review. Int J Med Inf 152(December 2020):104487. https://doi.org/10.1016/j.ijmedinf.2021.104487

Mabrouk A, Zagrouba E (2018) Abnormal behavior recognition for intelligent video surveillance systems: a review. Expert Syst Appl 91:480–491. https://doi.org/10.1016/j.eswa.2017.09.029

Nahavandi D, Alizadehsani R, Khosravi A, Acharya UR (2022) Application of artificial intelligence in wearable devices: opportunities and challenges. Comput Methods Programs Biomed 213(December). https://doi.org/10.1016/j.cmpb.2021.106541

Nekar DM, Kang HY, Yu JH (2022) Improvements of physical activity performance and motivation in adult men through augmented reality approach: a randomized controlled trial. J Environ Public Health 2022:3050424. https://doi.org/10.1155/2022/3050424

Ng YL, Ma F, Ho FK, Ip P, Fu KW (2019) Effectiveness of virtual and augmented reality-enhanced exercise on physical activity, psychological outcomes, and physical performance: a systematic review and meta-analysis of randomized controlled trials. Comput Hum Behav 99(September 2018):278–291. https://doi.org/10.1016/j.chb.2019.05.026

Ni MY, Hui RWH, Li TK, Tam AHM, Choy LLY, Ma KKW, Cheung F, Leung GM (2019) Augmented reality games as a new class of physical activity interventions? The impact of Pokémon Go use and gaming intensity on physical activity. Games Health J 8(1):1–6. https://doi.org/10.1089/g4h.2017.0181

Oh YJ, Zhang J, Fang ML, Fukuoka Y (2021) A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. Int J Behav Nutr Phys Act 18(1):1–25. https://doi.org/10.1186/s12966-021-01224-6

Ruiz-Ariza A, Grao-Cruces A, Marques NE, Martínez-López EJ (2017) Influence of physical fitness on cognitive and academic performance in adolescents: a systematic review from 2005–2015. Int Rev Sport Exerc Psychol 10(1):108–133. https://doi.org/10.1080/1750984X.2016.1184699

Ruiz-Ariza A, Casuso RA, Suarez-Manzano S, Martínez-López EJ (2018) Effect of augmented reality game Pokémon GO on cognitive performance and emotional intelligence in adolescent young. Comput Educ 116:49–63. https://doi.org/10.1016/j.compedu.2017.09.002

Smith B, Wightman L (2021) Promoting physical activity to disabled people: messengers, messages, guidelines and communication formats. Disabil Rehabil 43(24):3427–3431. https://doi.org/10.1080/09638288.2019.1679896

Soltani P, Morice AHP (2020) Augmented reality tools for sports education and training. Comput Educ 155(June 2019):103923. https://doi.org/10.1016/j.compedu.2020.103923

Szucs K, Brown E (2018) Rater reliability and construct validity of a mobile application for posture analysis. J Phys Ther Sci 30(1):31–36. https://doi.org/10.1589/jpts.30.31

Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN (2020) Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 8(7):1–13. https://doi.org/10.2196/15576

Wang Y, Cang S, Yu H (2019) A survey on wearable sensor modality centred human activity recognition in health care. Expert Syst Appl 137:167–190. https://doi.org/10.1016/j.eswa.2019.04.057

Wei S, Huang P, Li R, Liu Z, Zou Y (2021) Exploring the application of artificial intelligence in sports training: a case study approach. Complexity 2021. https://doi.org/10.1155/2021/4658937

West K, Purcell K, Haynes A, Taylor J, Hassett L, Sherrington C (2021) "people associate us with movement so it's an awesome opportunity": perspectives from physiotherapists on promoting physical activity, exercise and sport. Int J Environ Res Public Health 18(6):1–14. https://doi.org/10.3390/ijerph18062963

Winand M, Ng A, Byers T (2020) Pokémon "Go" but for how long? A qualitative analysis of motivation to play and sustainability of physical activity behaviour in young adults using mobile augmented reality. Managing Sport and Leisure 0(0):1–18

World Health Organization (2020) Physical activity. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/physical-activity. Accessed 5 Oct 2022

Chapter 15
Exergames, Artificial Intelligence and Augmented Reality: Connections to Body and Sensorial Experiences

Mateus David Finco, Vagner Ramos Dantas, and Vanide Alves dos Santos

Abstract This chapter introduces Augmented Reality (AR) as one of the most prominent technologies promoting immersive and sensorial experiences. The research aims to describe how Artificial Intelligence (AI) has been used through game engines in complex feedback systems (Deep Learning), as well as customizing efficient models to engage and monitor human interaction in an immersive way in activities such as health and well-being. It was possible to observe that different fields such as military training, education, cultural exhibition, and many others, are applying activities and developing many interactions mediated through AR. In addition, with the rise of wearables, it is clear there is a growing need to expand mechanisms that make the experience more realistic and interactive. One of the possible proposals would be the expansion of immersive environments with AR projection. Another growing area is entertainment, bringing new connections with Virtual Reality (VR) in the newest generation of exergames. The intention is to connect body movements with games through the sensorial immersion of AR, empowering new possibilities for the evolution of video games. This chapter presents a review of literature with the initial experiences involving AR and exergames through Deep Learning and to discuss the possible future of this emerging connection of sensoriality and body movements, supported by the AR tools and devices.

15.1 Introduction

The use of the most varied technologies has been advancing over time and lately it is possible to observe a special emphasis on the connections of body movements and sensorial experiences.

Moving from tactile hand commands through video game and computer joysticks, the body experiments also allowed different interactions between players and

M. D. Finco (✉) · V. R. Dantas · V. A. dos Santos
Psychopedagogy Department, Centre of Education, Federal University of Paraíba, João Pessoa, Brazil
e-mail: mateus.finco@academico.ufpb.br

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_15

machines: one could then observe the rise of exergames. Many experiences started in the 1970s, when the first video games also launched devices that allowed the amplification of body movements in different actions, such as balance platforms, mats with motion detectors and pistols for shooting (Finco and Maas 2014). However, the beginning of a trajectory of games using infrared sensors, took place from 2006 with the launch of Nintendo, the Wii video game. At that moment, the so-called exergames began.

The term "exergame" has been used in academia as well as in the industry, in slightly different ways, but usually it is agreed that these are active video games that combine body movement (exer = exercise) with gaming skills (games).

According to Finco et al. (2015) people are submerged in a society where technology pervades virtually all segments. In this context, digital games are one of the most widely used technological artifacts. Active video games, or exergames, were introduced as a way to allow the body to control such games.

Following the advancement of technologies, we can also see a great advance, based on the advent of exergames, of augmented reality (AR). The newest video games also have been investing in new devices, especially to raise the level of sensory immersion in games.

Augmented reality emerges as a new potential tool for the body interaction, especially for allowing varied experiences in the virtual world. According to Hayes (2020) augmented reality continues to develop and become more pervasive among a wide range of applications. Since its conception, marketers and technology firms have had to battle the perception that augmented reality is little more than a marketing tool. Also, the author stresses that AR's benefits can also extend to the healthcare sector, where it can play a much bigger role. For instance, AR apps enable users to see highly detailed, 3D images of different body systems when they hover their mobile device over a target image. This use of AR has become a powerful learning tool for training medical professionals (Hayes 2020).

The aim of this study is to analyse the initial experiences involving AR and exergames through deep learning and to discuss the possible future of this emerging connection of sensoriality and body movements, supported by the AR tools and devices.

15.2 Technologies and Learning

Technologies as learning tools are items that facilitate the learning process of individuals with more practicality and interactivity. In this sense, some technologies stand out and can contribute to learning, such as: Artificial Intelligence (AI), Augmented Reality (AR), Virtual Reality (VR), Robotics, etc. Some of these resources are already present in the routine of many people, which contributes to learning in different contexts. Present in several contexts, in addition to schools, it is possible to mention the use of AR in military training, as well as in cultural exhibition spaces, besides different educational and course training contexts.

In the military context, AR has been used to support strategies that aim to provide real-time location updates for hard to see targets, contributing to the optimization of soldiers' aiming performance, maintaining accuracy and efficiency under conditions of stress, load of work and uncertainty (Brunyé et al. 2020). In addition, AR was designed to provide the military with information specific to the environment in which they are inserted, and in turn, being relevant to the tasks, integrating with the equipment used in the missions, for example, implementing this technology in the glasses, in an intuitive and discrete, aiming to increase visual and multimodal perception, providing timely and actionable information (Azuma et al. 2001).

Regarding the context of cultural exhibition, AR has been implemented in digital reconstructions and their possibilities of visualization and interaction. Present in museums since the mid-2000s, enabling greater visitor interaction with the space and its works, such as: facilitating the route of visiting the museum, through adjacent resources and overlapping information; promote experiences beyond materiality, through insights into the history of cultural objects; access materials that are not visible due to security, logistics or that are undergoing maintenance; Digital reconstruction of cultural objects damaged by the action of time (Bozzelli et al. 2019).

According to Bozzelli et al. (2019) the use of AR in these environments allows for interactive and immersive experiences from applications of small contexts such as objects and museum windows, present for example in the Museum of Celtic Heritage. In addition to architectural contexts, present in Jumièges, Lumin and territory appreciated in Olympia and Carnuntum, allowing visitors a reconstructed view of the buildings during the tour along the old city from archaeological remains.

Some technologies such as VR and AI are being used more and more in an integrated way, to produce more immersive serious game environments, making the gameplay more attractive and dynamic, encouraging the player to stay longer, so that their game patterns data are recognized, and can contribute to the improvement of adaptive intelligent game systems. An example of this is the "Illumination Butterfly" project, a serious game that is based on neural network-driven 3DUI interaction techniques using two emerging machine learning algorithms: Generative Adversarial Imitation Learning (GAIL) and Proximal Policy Optimization (PPO) to learn exercises physically generated within the game (Elor and Kurniawan 2020).

15.3 Artificial Intelligence and Augmented Reality

The use of AI within AR applications has been discussed in different scenarios. Sahu et al. (2021) highlights that the integration of these technologies in production environments facilitates the exchange of information in a simplified way without changing the operator's attention, and is revolutionizing the industry, by incorporating intelligence within an automated infrastructure, making robotic systems autonomous, and able to make decisions based on their data.

Moawad et al. (2019) highlights that most AI models currently used have something in common, the ability to isolate repetitive trends and use them to accurately predict and direct future decisions. This is an important resource used in medicine to help improve patient care and ease the burden on the physician. The authors highlight that AR is currently being used in gynecological surgeries in conjunction with AI to guide surgeons in making correct decisions. Chen (2019) introduces an AR microscope with real-time AI integration for cancer diagnosis, and this has helped to standardize cancer assessment around the world, and expand healthcare technologies.

From a philosophical perspective, Chaudhary (2019) argues that AR is a disruptive technology, which produces an unprecedented continuous view, between reality and the digitally augmented medium, and the expansion of its tools, has been generating hardware and software ecosystems that has been pushing artists to develop features without requiring very complex programming. In this scenario, the "technologisation" of human perception and the experience of reality, combined with the development of AI-based natural language assistants, could lead to a secular re-enchantment of the world, with human existence being shaped through AR through digital avatars.

According to Norouzi et al. (2019), the emergence of exciting products such as Microsoft HoloLens; Metavision (Meta 2); Apple AR Developer Kits (ARKit); and well-funded startups like Magic Leap, is driving consumer interest, generating a return on developing this technology alongside others. This has been raising the expectation that soon AR together with AI such as intelligent virtual agents (IVAs) and the Internet of things (IoT) will be incorporated into our daily lives and will have a strong influence on our society in the near future.

Microsoft HoloLens has already inspired the development of applications that integrate various technologies in the field of education, such as the one by Roosan et al. (2022), called PGxKnow: an educational HoloLens application of AR and AI pharmacogenomics (PGx). It was developed using feedback from three clinical pharmacists trained in PGx and applied to 15 participants who independently reviewed the application and rated its usability using the System Usability Scale (SUS). It was evident that the use of this AR application for AI-based pharmacogenomics (PGx) education has the potential to fill the gap in PGx education, expand its use, and increase its impact in precision medicine.

Still in the context of education, Rajendran's (2019) proposal stands out: AREDAI Augmented Reality Based Educational Artificial Intelligence System, a multifunctional educational tool that can be widely used in an educational environment, aimed at this new generation of students in this environment current technology. The feature consists of an AR information display module with optional VR display, powered by image processing and optical character recognition. Through it, it is possible to detect predetermined objects and obtain information about them. In this tool there is an optical character recognition, capable of extracting text from books and obtaining additional information about it. Through the application it is possible to perform a secure search using keywords, blocking inappropriate content. In the gaming area, Keshav et al. (2019) developed Empowered Brain, a game based on AR and AI that assists in the socio-emotional communication of students with ASD. In this study, it is evident that the game also evaluates the student's performance, relating it to validated clinical measures of ADHD, suggesting that new technologies like this can produce new ways of identifying and characterizing ADHD symptoms in student populations. However, they claim that larger samples with more diverse participants are needed to validate these findings, although they agree that these technologies have the potential to help identify and longitudinally monitor such symptoms.

Training simulators in the health area was another topic addressed, which use AI, AR and Serious Games to improve the practical experience and performance of newly graduated professionals, to act in their respective attributions. An example of this can be observed through the study by Sitterding et al. (2019), in which a survey of the use of AI and games to improve the transition of new nursing professionals through VR and AR is presented. It lists tools such as the Neonatal Resuscitation Game and the AR resource CAE VimedixAR with Microsoft HoloLens Augmented Reality as a means of promoting practical action and cognitive stimulation. The study also highlights a greater effectiveness in learning through VR simulations than through the traditional method, and that the offer of AI, games and VR become an innovative approach capable of influencing the cognitive development of clinical reasoning.

In general, studies indicate that AI can contribute by making AR applications more dynamic, intelligent and personalized, adapting to different scenarios and situations, to promote a more intuitive and easier usability, helping in human development and its technologies, and has been expanding the perception of information about environments, objects and contents, adding detailed information, based on accurate analytical data, which go beyond an initial human perception, complementing us.

15.4 Exergames and Augmented Reality

Studies such as the one carried out by Willwacher and Korn (2021) have been presenting proposals that highlight the potential of gamification and exergames in the health area, through an intelligent framework for prevention and rehabilitation. Through it, it would be possible to carry out an autonomous, dynamic and personalized playful interactive training, through a hybrid system that combines the rules of experts with machine learning methods guided by data from the analysis of the performed movements.

Research such as the one developed by Jung et al. (2020) sought to understand in practice, in a longitudinal way, how this dynamic would be established between therapists and stroke patients within a routine of game-assisted therapy in a clinical setting. Considering that gamification and games proved to be resources with the potential to make tedious and repetitive rehabilitation activities more fun, increasing people's motivation to perform these tasks, making them faster, more frequent and pleasurable.

Throughout the research, Jung et al. (2020) noticed different patterns of engagement in game-assisted therapy, indicating a need to adapt games to the specific needs of each patient, to increase their involvement in a therapeutically meaningful way.

And it highlighted the need to develop user interfaces dedicated to training, to broaden the support of therapists during sessions.

Facts like these demonstrate the need for investment in game design development in clinical environments. Considering that stimulating personalised interfaces, and with detailed feedback systems, added to options for modulating the difficulties of activities within playful immersive virtual reality environments, have been promoting greater engagement in autonomy while performing tasks, and increasing the chances for new tasks to be performed (Elor and Kurniawan 2020).

From the analysis of the insertion of gamification and exergames for monitoring movements in the training for prevention and rehabilitation of patients, an increased risk of injuries was noticed, demonstrating the need for dynamic adjustments in the difficulty at the patient's level in real time. The increase in risk may be a consequence of increased motivation, which generates greater immersion in the game, which may reduce the perception of stimuli, such as pain, and increase the risk of falling (Willwacher and Korn 2021).

The advancement of gamification techniques and the variety of games, with their increasingly diverse elements, have been stimulating the service to different profiles of player users. Given this demand, new possibilities for games emerged, with the advancement of technological devices with sensors, enabling the practice of more elaborate exercises (Willwacher and Korn 2021). The author also stated that with the advent of game consoles using sensor-based motion control such as the Nintendo Wii and Microsoft Kinect, it became possible to capture physical activities playfully with comparatively little technical effort.

In addition to commercial titles for home training such as Wii Sports (Nintendo 2006), Kinect Sports (Microsoft 2010), or Shape up (Ubisoft 2014), games based on motion sensors were also developed specifically for prevention or rehabilitation (see Fig. 15.1). Such training games are called "exergames", a portmanteau of the words "exercise" and "game" (Willwacher and Korn 2021, p. 856).

15.5 Reinforcement Learning

According to Recht (2019) Reinforcement Learning (RL) is the subfield of machine learning that studies how to use past data to enhance the future manipulation of a dynamical system. A control engineer might be puzzled by such a definition and interject that this is precisely the scope of control theory. That the RL and control communities remain practically disjoint has led to the co-development of vastly different approaches to the same problems. The author also states that a tremendous opportunity lies in deploying its data-driven systems in more demanding interactive tasks, including self-driving vehicles, distributed sensor networks, and agile robotic systems. For RL to expand into such technologies, however, the methods must be both safe and reliable—the failure of such systems can have severe societal and economic consequences, including the loss of human life.

Fig. 15.1 The game shape up. Cover from the game "Shape Up". Microsoft. 10 January 2023. www.xbox.com/en-US/games/store/shape-up/c1b6dl0t68q5

François-Lavet et al. (2018) describe how the RL issue can be structured through an agent that needs to make decisions in an environment to optimize its notion of cumulative rewards. This process can be broadly applied to a variety of tasks and captures many essential AI features, such as a sense of cause and effect, and those of uncertainty and non-determinism. And they point out that a fundamental aspect of RL is that an agent must be able to learn good behavior, through its modifications and the acquisition of new behaviors and skills in an incremental way, such as the experience of trial and error. Thus, the authors emphasize that the RL agent does not require complete knowledge or control of the environment, it only needs to be able to interact with the environment and collect information.

By allowing computers to learn from their own experience (machine learning), understanding the world in terms of a hierarchy of concepts, complex graphs with many layers are generated, giving rise to the approach known as AI deep learning (Goodfellow 2016). It is a subdivision of AI capable of making precise decisions based on data, allowing computational models composed of several layers of processing, capable of learning the representations of the data, from multiple levels of abstraction. This allows the discovery of complex structures in large datasets using the backpropagation algorithm (LeCun et al. 2015).

Mousavi et al. (2016) highlights that this machine learning method called deep learning has been gaining a lot of attention, due to its surprising results in several areas in which it has been applied, such as pattern recognition, speech recognition, computer vision and data processing natural language. And recently, deep learning techniques have been combined with reinforcement learning methods, which involve learning strategies through interaction with the environment, to expand problem solving, culminating in a new research route called deep reinforcement learning.

RL is described by Shrestha (2019) as being something specifically suited to problems involving short- and long-term rewards, as is the case of the game of Go, such as AlphaGo, the Google program that was able to beat the champion of Human Go through RL learning. They point out that when combining deep network architecture with RL, DRL is generated, which can extend the use of reinforcement to games and more complex areas such as robotics, healthcare and finance. With the advent of DRL, problems that seemed unsolvable with RL can now be solved by increasing the number of hidden layers of deep networks and RL-based algorithms that maximize the reward for the actions taken by the agent.

Currently, works such as by Elor and Kurniawan (2020) have been using deep reinforcement learning by integrating game engine systems (Unity Game Engine); ML-Agents together with serious games, such as immersive virtual reality exergames with the HTC Vive Head-Mounted Display, to generate machine learning capable of creating promising predictive models, to control the system and understand user behaviour. Specifically, this technology utilized neural network agents to enhance the gameplay experience, in which a virtual robot arm assisted the user in exercising to protect butterflies from projectiles.

Brockman et al. (2016) present a discussion about the components of OpenAI Gym and the design decisions that were included in this software, which is a toolkit for research with RL. In it can be seen a growing collection of benchmark problems that expose a common interface and a website where people can share their results and compare the performance of algorithms. Ray (2019) emphasizes that safety with AI is a critical concern, and some errors become unacceptable, such as robotics systems that interact with humans and that should never cause injury to them during exploration. In this sense, they are currently training AI agents via RL, where security concerns are minimal, anticipating real-world challenges and complexities through simulation. One solution to this is the OpenAI Safety Gym benchmark suite, a safe place for AIs to learn, through environments under continuous high-dimensional control, in which the progress of research in RL is forcibly measured.

15.6 Sensorial Experience and Artificial Intelligence

The search for the development of more accurate immersive experiences, such as the therapeutic use in prevention and rehabilitation, has included the use of movement sensors and perception of the human condition, through affective computing and sentiment analysis. Through them, it is possible to interpret the human emotional state and adapt their behaviour according to their needs and potential, personalizing activities, through the recognition of certain patterns, through data collected through machine learning (Cambria et al. 2017). Information processing takes place in different modalities such as: speech recognition, natural language processing and facial expression detection.

Current developments in wearable sensor technology, e.g., using “Inertial Measurement Units” (IMUs) and machine learning techniques, allow quantifying the movement and load characteristics of an exercise without therapeutic personnel. These systems are currently developed mainly for use in gait and running analyses, but in principle, this method can be extended to typical prevention and rehabilitation exercises (Willwacher and Korn 2021 p. 857).

The need for accurate and real-time monitoring of human exercises, in the absence of a professional therapist, has expanded the type and variety of sensors that can provide such information. This has been driving new technologies with increasingly complex AI, which are intended to capture information through various sensors, such as movement capture devices, and information resulting from the metabolites of the developed exercises, to customize new activities adapted to the needs and possibilities of the patient.

The exercise is monitored using smart sensing approaches. Depending on the type of exercise, different sensors can be applied: sensors worn on the body (i.e., IMUs in wristbands, belts, or implemented in smart apparel) or cameras can track the person's motion. Further, the concentration of specific metabolites (i.e., lactate) through an analysis of the produced sweat on the skin could be monitored and considered in evaluating the intensity of the exercise for the patient. An AI-supported analysis of facial expressions could further determine the subjective intensity (Willwacher and Korn 2021 p. 858).

Through recent studies (Elor and Kurniawan 2020; Willwacher and Korn 2021) there is a concern with precision in capturing and analysing human movement data, through an automated and increasingly complex AI, to produce exercises with an adapted level of difficulty to the possibilities and interests of the patient. This requires a level of detail through a system of rules developed by experienced professionals (experts) from different areas, to produce through AI, a decision tree or a machine learning system.

Willwacher and Korn (2021) states that adequate automated adaptation of the difficulty of training measures requires pattern analysis and pattern recognition beyond pure motion detection, i.e., collecting raw kinematic data and measuring the performance progression in detail concerning what is happening in the game. Thus, a specific game situation can justify an increased performance level in the short term without increasing the overall difficulty. For this purpose, a system of rules must be established, which might be best formalized via a classical element of AI, the decision tree. Such rule-based systems have also been established in the healthcare sector for some time as "rule-based decision support systems" or simply as medical decision aids. In principle, experts (e.g., sports scientists in collaboration with physicians and computer scientists) can either develop decision trees or automatically induce them from data using machine learning techniques. There are several competing algorithms for the latter, particularly CHAIDs (Chi-square Automatic Interaction Detectors) and CARTs (Classification and Regression Trees). Currently, the most promising solutions for such training in healthcare are probably so-called hybrid systems: systems combining rules from human experts with data-driven machine learning methods for movement analysis (Willwacher and Korn 2021, p. 859).

Therefore, the literature has been pointing out (Willwacher and Korn 2021; Jung et al. 2020; Elor and Kurniawan 2020) that the likely most promising solutions today in training for health promotion are developed through hybrid systems, through engagement with the use of immersive games in repetitive and dynamic exercises, combining rules established by experienced professionals and machine learning methods driven by the data of movement analysis.

15.7 Conclusion

Body and sensorial experiences through the use of technologies are growing and it is possible to observe a large expanding market involving AI and AR. The use of these innovative resources can enable new experiences that will directly favor the development of several areas, from specific job training, education and especially in the health area.

Exergames, since the beginning of their appearance in the technological market, have proposed devices that amplify large body movements and with that, provide new sensory connections and motivation both in terms of learning and entertainment. From this point of view, the contribution of AI can also be seen so that the machines are autonomous in recognizing the users' needs for the proposed interaction and objectives.

Besides the rise of wearables, it is clear that there is also a growing need to expand mechanisms that make the experience more realistic and interactive, and one of the possible proposals would be immersion environments with AR projection, which would allow for example movement with greater naturalness and without the use of implements, characterizing the human experience.

For future studies, it becomes relevant to analyze the new advances it becomes essential to evaluate the effectiveness of AI systems in the interaction of activities developed in AR environments and in the observation of the potential that exergames can develop in the use of such technologies.

References

Azuma R, Baillot Y, Behringer R et al (2001) Recent advances in augmented reality. IEEE Comp Graph App 21(6):34–47. https://doi.org/10.1109/38.963459

Bozzelli G et al (2019) An integrated VR/AR framework for user-centric interactive experience of cultural heritage: The ArkaeVision project. Dig App Arch Cult Herit e00124. https://doi.org/10.1016/j.daach.2019.e00124

Brockman G et al (2016) Openai gym. ArXiv. https://doi.org/10.48550/arXiv.160601540

Brunyé TT, Brou R, Doty TJ et al (2020) A review of US army research contributing to cognitive enhancement in military contexts. J Cogn Enhanc 4(1):453–546. https://doi.org/10.1007/s41465-020-00167-3

Chen P et al (2019) An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. Nat Med 25(9):1453–1457. https://doi.org/10.1038/s41591-019-0539-7

Cambria E et al (2017) Affective computing and sentiment analysis. In: Cambria E, Das D, Bandyopadhyay S, Feraco A (eds) A practical guide to sentiment analysis. Socio-affective computing, vol 5, pp 1–10. https://doi.org/10.1007/978-3-319-55394-8_1

Chaudhary MY (2019) Augmented reality, artificial intelligence and the re-enchantment of the world: with Mohammad Yaqub Chaudhary,"Augmented Reality, Artificial Intelligence, and the Re-Enchantment of the World"; and William Young,"Reverend Robot: Automation and Clergy." Zygon® 54(2):454–478. https://doi.org/10.1111/zygo.12521

Elor A, Kurniawan S (2020) Deep reinforcement learning in immersive virtual reality exergame for agent movement guidance. In: 2020 IEEE 8th international conference on serious games and applications for health (SeGAH) IEEE, pp 1–7. https://doi.org/10.1109/SeGAH49190.2020.9201901

Finco MD, Maas RW (2014) The history of exergames: promotion of exercise and active living through body interaction. In: Proceedings of the 3rd international conference on serious games and applications for health (SeGAH), Federal University Fluminense, Rio de Janeiro, June 2014, pp 14–16

Finco MD et al (2015) Exergames laboratory: a complementary space for physical education classes. Movimento J 21(3):687–700. https://doi.org/10.22456/1982-8918.52435

François-Lavet V et al (2018) An Introduction to deep reinforcement learning. Found Trends® Mach Learn 11(3–4):219–354. http://dx.doi.org/https://doi.org/10.1561/2200000071

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press. <https://www.deeplearningbook.org/>

Hayes A (2020) Augmented reality (AR) defined, with examples and uses. https://www.investopedia.com/terms/a/augmented-reality.asp. Accessed 10 Sep 2022

Jung HT et al (2020) Rehabilitation games in real-world clinical settings: practices, challenges, and opportunities. ACM Trans Comput-Hum Inter (TOCHI) 27(6):1–43. https://doi.org/10.1145/3418197

Keshav NU et al (2019) Digital attention-related augmented-reality game: significant correlation between student game performance and validated clinical measures of attention-deficit/hyperactivity disorder (ADHD). Children 6(6):72–83. https://doi.org/10.3390/children6060072

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

Moawad G, Tyan P, Louie M (2019) Artificial intelligence and augmented reality in gynecology. Cur Op Obst Gynecology 31(5):345–348. https://doi.org/10.1097/GCO.0000000000000559

Mousavi S S, Schukat M, Howley E (2016) Deep reinforcement learning: an overview. In: Bi Y, Kapoor S, Bhatia R (eds) Proceedings of SAI intelligent systems conference (IntelliSys) 2016. IntelliSys 2016 (Lecture notes in networks and systems), vol 16, pp 426–440. https://doi.org/10.1007/978-3-319-56991-8_32

Norouzi N et al. (2019) A systematic review of the convergence of augmented reality, intelligent virtual agents, and the internet of things. In: Al-Turjman F (eds) Artificial intelligence in IoT. Transactions on computational science and computational intelligence. Springer, Cham, pp 1–24. https://doi.org/10.1007/978-3-030-04110-6_1

Rajendran PS, Christian IS, Shedge MS (2019) AREDAI augmented reality based educational artificial intelligence system. Inter J Rec Tech Eng (IJRTE) 8(1):1960–1964

Ray A, Achiam J, Amodei D (2019) Benchmarking safe exploration in deep reinforcement learning 7(1). arXiv preprint arXiv:1910.01708

Recht B (2019) A tour of reinforcement learning: the view from continuous control. Annu Rev Control Robot Auton Syst 1(2):253–279. https://doi.org/10.1146/annurev-control-053018-023825

Roosan D et al (2022) PGxKnow: a pharmacogenomics educational HoloLens application of augmented reality and artificial intelligence. Pharmacogenomics 23(4):235–245. https://doi.org/10.2217/pgs-2021-0120

Sahu CK, Young C, Rai R (2021) Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review. Int J Prod Res 59(16):4903–4959. https://doi.org/10.1080/00207543.2020.1859636
Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. IEEE Access 7:53040–53065. https://doi.org/10.1109/ACCESS.2019.2912200
Sitterding MC et al (2019) Using artificial intelligence and gaming to improve new nurse transition. Nurse Lead 17(2):125–130
Willwacher S, Korn O (2021) Gamification of movement exercises in rehabilitation and prevention: a framework for smart training in AI-based exergames. In: Shin CS, Di Bucchianico G, Fukuda S, Ghim YG, Montagna G, Carvalho C (eds) Advances in industrial design. AHFE (Lecture notes in networks and systems,vol 260, pp 855–862. https://doi.org/10.1007/978-3-030-80829-7_104

Part IV
Combining Augmented Reality and Artificial Intelligence to Enhance Services, Retail, and Recommendations

Chapter 16
FUSE: Towards AI-Based Future Services for Generating Augmented Reality Experiences

Klen Čopič Pucihar, Vladimir Geroimenko, and Matjaž Kljun

Abstract This chapter explores the use of Artificial Intelligence (AI) to semantically understand the scene, to decide when to provide Augmented Reality (AR) content or to take other actions, and how to generate the augmented content. The chapter presents a survey of novel Machine Learning techniques that have a capability to assist future AI-driven smart AR services and to enable their generalisation to a wide range of scenarios making them ubiquitous, adaptable and robust for real-world usage. The key challenge considered by the authors is how to generalise these techniques to unprepared arbitrary environments in which the scene objects are automatically detected, recognised and segmented in an intelligent way that allows the system to annotate the scene automatically. The chapter begins with an exploration of various methods and techniques which could enable the creation of AR services at such a high level of generalisation. It provides a review of a variety of machine learning algorithms for object detection, recognition and image segmentation. Then, it analyses 3D scene creation and reconstruction, which includes various 3D scanning approaches as well as techniques for generating context aware text and image annotations. Finally, the chapter considers the newest methods of view synthesis using neural networks known as 3D neural rendering and finishes with a discussion about the current and future challenges of AI-based services for generating AR experiences.

K. Č. Pucihar (✉) · M. Kljun
Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia
e-mail: klen.copic@famnit.upr.si

M. Kljun
e-mail: matjaz.kljun@upr.si

Faculty of Information Studies (FIŠ), Novo Mesto, Slovenia

V. Geroimenko
Faculty of Informatics and Computer Science, The British University in Egypt, Cairo, Egypt
e-mail: vladimir.geroimenko@bue.edu.eg

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_16

16.1 Introduction

One of the key challenges of creating smart AR-based services is to make sure they can sufficiently generalise. For example, an AR system for translating texts of a foreign language has to be able to translate non-predefined text printed on various types of surfaces (e.g., traffic signs, displays, paper, product packaging) using various typefaces and their sizes. Or a smart digital camera, which automatically determines when to take the best picture such as when everyone in the picture is smiling, should be able to do so without any prior knowledge on the number of people, their ethnicity, age, or any other information regarding their appearance.

AI-enhanced AR services show great potential for supporting learning in various contexts, including specialised training and language learning. For example, one of the popular techniques to learn vocabulary in a foreign language is the so-called 'keyword method', in which people use what a foreign word sounds like in their native language and create a memorable visual link between both words (e.g., a Japanese word for "book" is "hon", which sounds like "honey" and the memorable link can be "a book covered in honey"). In such a context, AR systems can annotate the real-world scene with memorable links in a foreign language (see Fig. 16.1). However, as prior results showed, this method can be further improved if visualisation (image) of the memorable link is also presented (Weerasinghe et al. 2022). Again, the key challenge is how to generalise this technique to unprepared arbitrary environments in which the scene objects are automatically detected, recognised and segmented so that the system can automatically annotate the scene. Furthermore, the system needs to determine appropriate sets of memorable links and automatically generate and place synthetic images around physical objects to aid in memorisation of words in a foreign language.

For an AR application to succeed in generating believable general experiences, it needs to implement the rendering pipeline depicted in Fig. 16.2. At the beginning of the pipeline is an optional set of sensors that capture information about the environment. These sensors may include one or more RGB cameras, depth cameras, IMU (inertial measurement unit), IR (infra-red) cameras to name a few. The information from the sensors is then processed with two goals in mind: (1) to calculate camera extrinsic parameters (parameters that define the camera pose such as position and orientation) and (2) to build an understanding of the environment that will enable the generation of augmentations in the form of text, images, 3D objects (defined as polygon meshes), or even neural scene representations. Once the extrinsic parameters together with the augmentations are ready, we can feed them to the renderer in order to generate 2D images that can be displayed on devices, such as a mobile phone or a head-mounted display. The calculation of extrinsic parameters commonly runs at every frame, whilst building understanding of the environment commonly runs on a request/response basis.

Defining extrinsic camera parameters is a well understood problem, which has been extensively studied (Qi et al. 2010). Despite several shortcomings, the systems

Fig. 16.1 The user is wearing a Head-Mounted Display (HMD) and observing the desk with different objects positioned on it. Utilising object recognition, identification and segmentation the system annotates the scene with virtual buttons. When pressed, the system reveals a keyword together with the English and Japanese words. The system also reads aloud the correct pronunciation of the word in Japanese and shows a memorable link "Hug a key" (Weerasinghe et al. 2022)

are generally very good in performing camera pose tracking in unprepared environments. As such, camera pose tracking can generalise reasonably well to various real-world scenarios. Contrary is true for methods that generate augmentations. Within research literature and commercially available applications these are commonly limited to predefined environments and rarely enable good generalisation that scales beyond several prototype scenes. This is because such a ubiquitous operation that generalises to a wider array of unprepared scenarios is far from trivial. However, it also presents an essential driver and an opportunity for the uptake of future AI driven services in the AR domain.

In this chapter, we consider various methods and techniques which could enable the creation of AR services at such a high level of generalisation. We provide a review of a variety of machine learning algorithms for object detection, recognition and image segmentation. Then, we analyse 3D scene creation and reconstruction, which includes various 3D scanning approaches as well as techniques for generating context aware text and image annotations. We also look at the newest methods of view synthesis using neural networks known as 3D neural rendering. The chapter finishes with a discussion of future services and challenges.

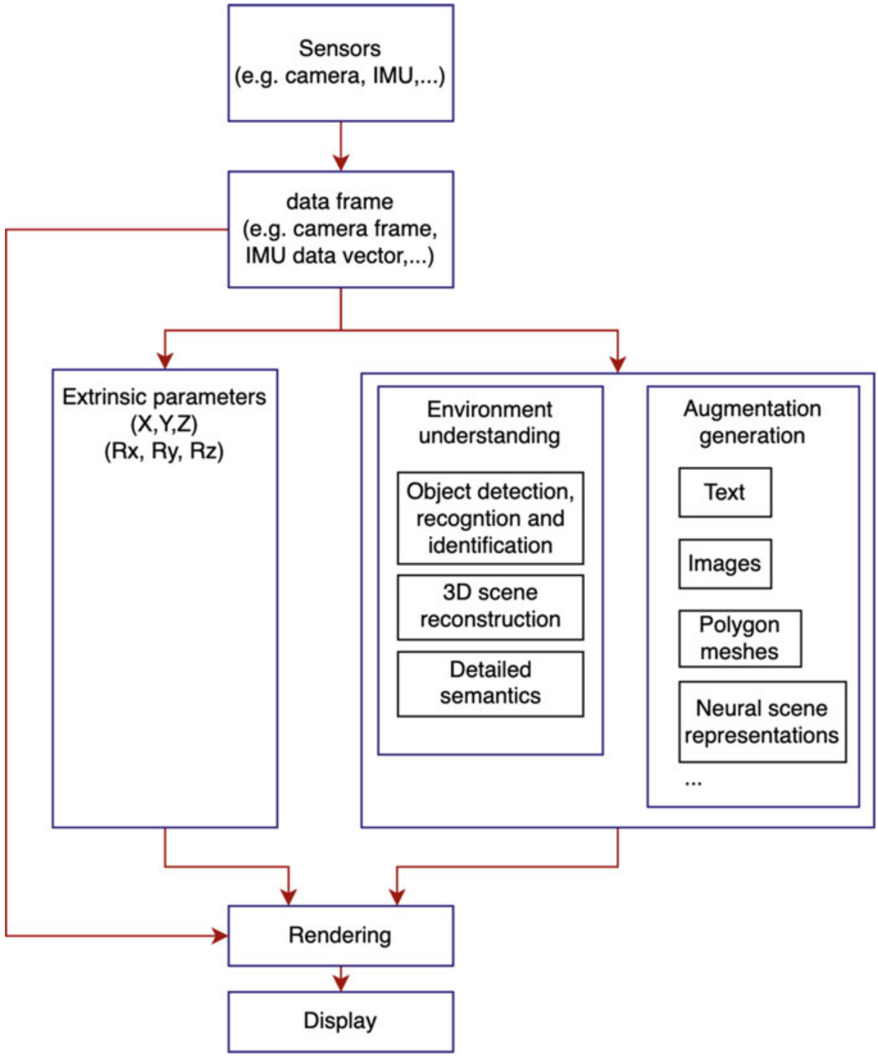


Fig. 16.2 AR rendering pipeline. A set of sensors capture information about the environment. The information is processed to: (**1**) calculate camera extrinsic parameters and (**2**) build understanding of the environment in order to generate augmentations in the form of text, images, 3D objects or neural scene representations. Finally, the extrinsic and graphic primitives are fed to the renderer which creates 2D images ready to be displayed

16.2 Understanding the Environment

In this section, we consider various techniques for the system to understand the environment. This includes detecting objects in the view of the user, understanding the environment such as for example understanding individual buildings on a town square, and providing additional semantic knowledge to the information obtained.

16.2.1 Object Detection, Recognition and Segmentation

Detecting and recognising objects in the scene is a vital component of AR applications. This is usually accomplished on 2D image feeds that come from one or more cameras. However, in object detection only coarse segmentation is performed, which can in some cases limit the dexterity of scene understanding (e.g., making it difficult to correctly highlight or annotate a 3D scene). This can be done via semantic segmentation. Figure 16.3 shows how the output of object detection draws a rough border around detected objects, semantic segmentation provides the exact outline of where in the scene the objects are, whilst the instance segmentation combines both to precisely recognise each instance of an object within the image.

If such detection, recognition and segmentation can generalise well over a variety of different real-world scenarios (e.g., lighting conditions, different environments, number of objects and classes), they offer valuable information for generating meaningful generalisable AR applications. However, such information is not by itself registered in 3D space, but, if rendered through a video see-through system (looking at the scene of the world through a video feed of the device camera) it enables several interesting AR applications such as automatic annotation generation. As already mentioned in the introduction, one interesting use of such a technology is AR support language learning in which the objects that surround the user are annotated/highlighted with labels in a foreign language (Ibrahim et al. 2018; Weerasinghe et al. 2022).

The recognition can also focus on identifying the activity within the scene. Such information could be useful, for example, to trigger support within the augmented view of the user. In the case of a smart digital camera that "knows" when to take a picture, detection relates to identifying peoples' faces whilst the recognition would relate to comprehending facial expressions (e.g., are people within the scene smiling or not). Or if an AR system is to translate the text in view of the camera, it first needs to identify whether there is any text in the scene (i.e., object detection), and once the text is identified it needs to read/identify each letter of the text (i.e., recognition). Only after this stage can the system attempt to translate the text and re-project it back to the scene. In order to accomplish the latter, it needs to understand the relative position of the camera in relation to the plane where the text is to be re-projected.

To date, several robust object recognition methods exist and can be divided into non-neural network approaches and neural network approaches. In both cases, the

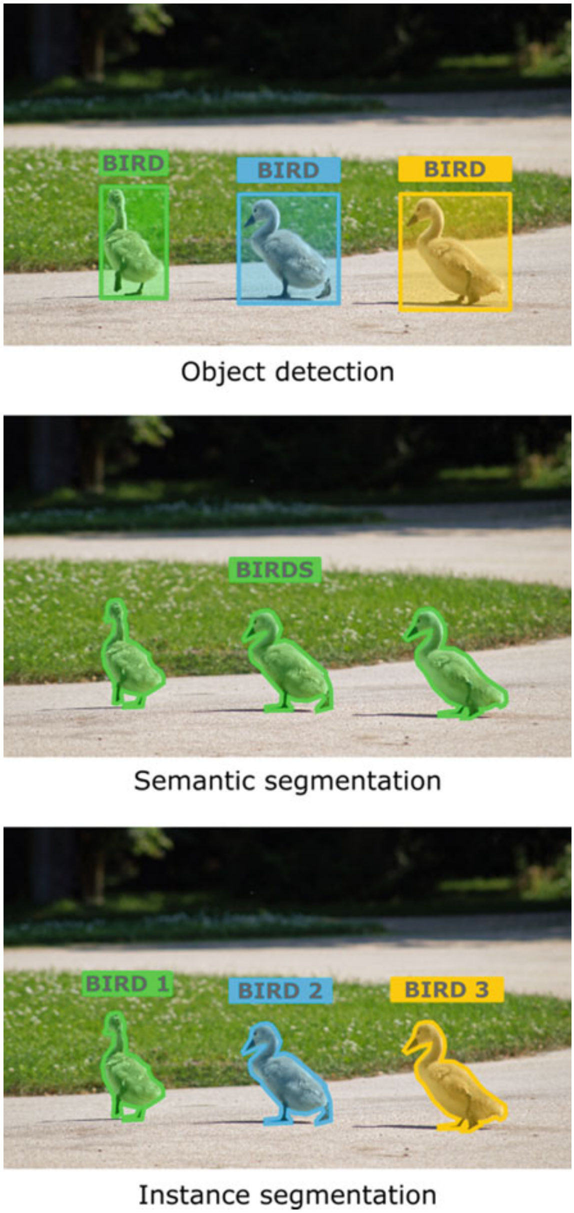


Fig. 16.3 Object detection, semantic segmentation, and the combination of both resulting in instance segmentation. The default image is the courtesy of last_firstborn from Pixabay

concept is very similar. The system attempts to define classes of objects by extracting distinct features for each object class. How such features are extracted and used is the main difference amongst practical applications we see. In non-neural network approaches we first have to come up with a set of features that describe classes (e.g., histogram of oriented gradients also known as HOG (Dalal and Triggs 2005),

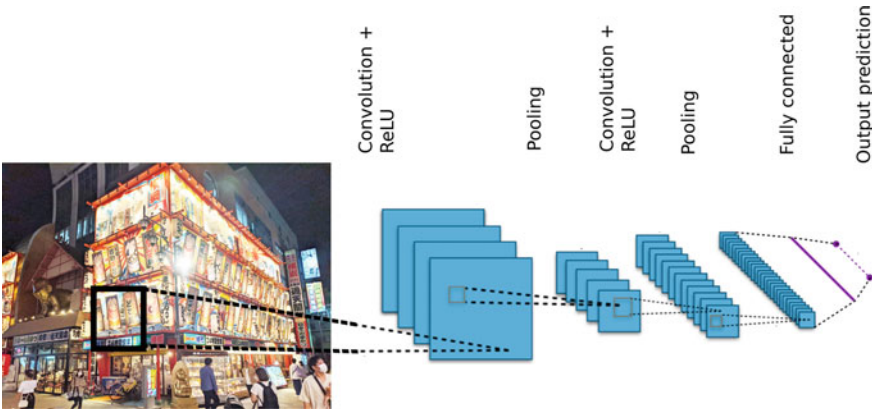


Fig. 16.4 An example of an object recognition system following a convolutional neural network approach. The model is learned through supervised learning on a training data set[1]

scale-invariant feature transform also known as SIFT (Lowe 2004)) and only after this is known can we run classification on regions of interest (e.g., random forest, logical regression, colour histograms) in order to detect objects. In the case of neural network approaches both feature extraction and classification are commonly handled by the neural network. The most common networks to be used are convolutional neural networks (CNN). Figure 16.4 visualises a common architecture of a CNN framework and is summarised based on (Ren and Wang 2022).

16.2.2 3D Scene Reconstruction

In addition to identifying objects in the scene, AR applications can also benefit from knowledge of the scene's geometry. The AR system needs to know its relative position in relation to the 3D scene. In a minimalist case, the scene can consist only of a single plane on which the augmentations are rendered in a geometrically coherent way (e.g., ground plane). However, such simplistic 3D scenes usually hold only a limited amount of required information about the environment. For example, they are not suitable for training applications where complex geometries are at hand, such as the car engine repair training scenario, which is depicted in Fig. 16.5. Here the repair instructions are overlaid on top of a car engine, which has a complex geometry; thus, the planar approximation for overlaying spatially complex instructions will not suffice. Further still, there are AR applications scenarios in which one would like to support or train users within large scale environments, such as training construction engineers when reviewing the construction site progress. As such the reconstruction

[1] Image adapted from Wikimedia Commons, the free media repository. Original authored by Aphex34. https://commons.wikimedia.org/wiki/File:Typical_cnn.png.

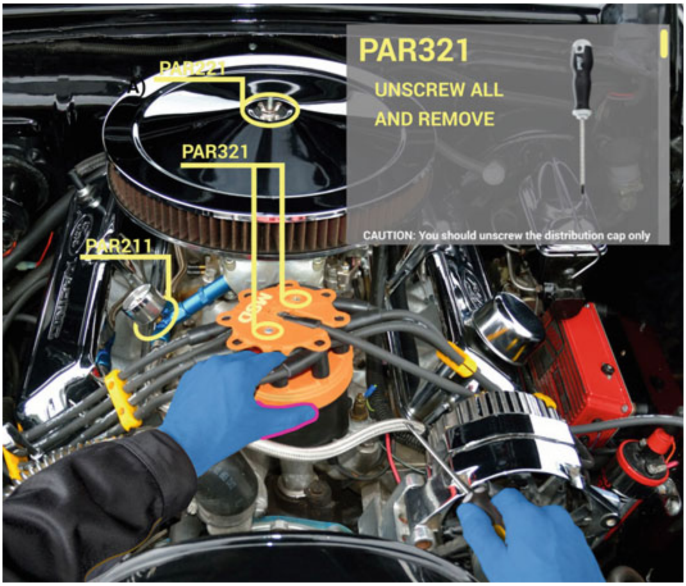


Fig. 16.5 An illustration of an augmented reality application in which instructions are provided to the mechanic in the form of textual annotations overlaid on top of a complex geometric object of a car engine. The instructions are only visible when wearing a head mounted device

of complex geometries and large-scale environments is important also from the point of extracting relevant semantic information about the environment. This can be done following several 3D reconstruction methods ranging from 3D scanning using laser scanners to structured light and stereoscopic cameras. Another interesting approach is to use RGB cameras and photogrammetry. Hereafter we review these techniques.

16.2.2.1 Photogrammetry

Photogrammetry refers to obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images (Mikhail et al. 2001). One application of photogrammetry is creating 3D models, which involves using 3D scanning technology to create digital copies of real-world objects such as manufactured parts and assemblies, free-form models made of clay, human anatomy and environmental scenes. 3D scanning is a process that involves taking multiple photos of an object, a room or environments from different angles. These photos are then analysed by software that creates a point cloud and mesh of the object, sometimes together with a texture derived from the photos. This method can produce high quality results, but it requires a good camera, preferably a laser scanner, a powerful computer, and specialised software (Pepe and Costantino 2020). However, there are also budget options available, such as photogrammetry programs that can be used on modern smartphones and a desktop computer (Collins et al. 2019).

Photogrammetry is today used in a variety of fields such as topographic mapping (Bi et al. 2017), architecture, engineering, manufacturing, geology (Bemis et al. 2014; Tamiminia et al. 2020), archaeology, and cultural heritage (Hall et al. 2001). Especially in the latter, accurately creating 3D models of heritage and archaeological objects and sites in their current state is essential for preserving our natural, cultural, and mixed heritage, which is constantly under threat from various sources such as military conflicts, natural disasters, climate change, and human neglect. The use of range sensors and imaging devices has been particularly helpful in documenting and preserving the built environment and natural heritage (Remondino 2011).

It has to be taken into account that achieving good results with this method requires a lot of effort, including taking hundreds of photos from different angles and potentially waiting for hours for the model to compile. In addition, since photogrammetry relies on images, dark, shiny or clear surfaces of an object being scanned can result in gaps or holes in the mesh, that require post-process clean-up with software like MeshLab, netfabb or MeshMixer (Enesi et al. 2022). As such, scene reconstruction with this approach needs to be performed in advance and if the scene in the AR scene changes, the geometry (and lighting) of the model differs from the real scene. Nevertheless, the reconstructed environment and objects within provide the AR system with data and geometry about the 3D scene looked at. In addition, the 3D scene can have semantic information about objects in the scene. For example, if the model of the town square is provided to the AR system, when the person is on the square mentioned the AR system understands the objects in the scene from the model.

16.2.2.2 3D Scanning Using Structured Light Sensors, Laser Sensors or Stereoscopic Images

3D scanning is possible also though precise depth maps which can be obtained by industrial computed tomography scanners (Fried et al. 2020), structured-light 3D scanners (Rocchini et al. 2001), LiDAR (Wulder et al. 2012) and Time Of Flight 3D scanners (Cui et al. 2010). 3D scanning is the process of collecting data on the shape (depth maps) and possibly the appearance (e.g., colour textures) of the real-world objects or environment. The 3D point clouds obtained with these capturing devices can then be numerically approximated to generate polygon meshes and voxel grids that represent object surfaces and volumetric structures. They can also be coupled with the RGB camera feed to capture in detail the appearance of textured surfaces (Fried et al. 2020). Such approaches can produce good results; however, they still require one to capture data from all object angles, which can be time consuming. Inferring “unseen” parts of the object can be solved with probabilistic guessing about the object’s appearance.

Nevertheless, 3D scanning presents a viable alternative to aforementioned approaches to provide 3D maps that can be used for camera pose tracking, and to reconstruct geometrically precise 3D models of the physical scene either preprocessed or in real time (Izadi et al. 2011). Providing geometry-aware information to augmented reality systems in real time is beneficial if the scene is changing. For example, understanding where planar or non-planar physical surfaces are located in an ever-changing geometric scene (e.g., disassembling the engine presented in Fig. 16.5), is very important for the AR system that appropriates the reconstructed surfaces for visualising augmented content as well as for interaction.

16.2.3 Detailed Semantic Understanding of the Environment

Semantic understanding of the environment could go beyond detecting, recognising and segmenting objects within a reconstructed 3D environment. For example, additional information about objects themselves could be also included. Such semantic information could be added utilising object appearance and localised queries scraping information from the World Wide Web. This information could include semantic data that is relevant for a particular AR use case. For example, a tourist application could query for semantic data about object name, its creator, purpose, a brief history surrounding the object and perhaps even a story of how it came to be in this particular location. Further still, it could even hold other dynamic properties about the object, such as object popularity and condition.

An alternative to such predefined semantic data caching is a system that generates human-like text responses in the form of a dialog. This could enable a conversion-like paradigm where the surrounding context plays a vital role in it. For example, based on the object at hand several predefined questions could be offered to the user, or the user could pose their own question. These could even be personalised as is common in case of recommendation systems. Based on the input data, the system would provide the response initiating a conversation. This paradigm is becoming ever more possible with recent advances in autoregressive AI language models that are based on deep learning such as ChatGPT[2] (released to the public in November 2022) and have greatly advanced the field of natural language processing (Brown et al. 2020).

16.3 Generating Augmentations for AR Applications

AR applications can augment real-world experience in visual, audio, haptic and olfactory modalities. In this chapter, we focus on visual augmentations, which are based on rendering graphics in a geometrically coherent way within the real-world environment, where augmentations can be observed through a video or optical see-through display (looking at the scene of the world through a transparent display).

[2] ChatGPT https://chat.openai.com/.

Most commonly, 3D graphics are rendered based on explicit or implicit representations of a 3D scene where the geometry and material properties are defined. These representations are then rendered via forward rendering approaches (e.g., rasterization) or ray casting respectively. However, manual generation of these scene representations is difficult and time consuming. For instance, the most common scene representation uses explicit surface definitions, such as, for example, polygon meshes composed of triangles or quads which are defined through a series of vertices. To control visual appearance of these explicit representations we also define visual properties, such as texture maps and other material properties. Another type of representation is volumetric representation in a form of voxels, which can store density values representing opacity and visual appearance of the scene. However, as already mentioned, generating optimised explicit meshes and implicit voxel representations is difficult and time consuming as they are either created by 3D artists or via 3D scanning or inverse rendering techniques. An alternative to this deterministic approach is to try and create these scene representations using a combination of computer graphics and machine learning algorithms, a method known as neural rendering. Hereafter follow subsections which will look at various methods for content generation, ranging from 3D modelling, and 3D scanning to neural rendering techniques.

16.3.1 3D Modelling

Two of the techniques described earlier, namely photogrammetry and 3D scanning, can be used to construct digital 3D models. Another common approach to create the content to be used as augmentations in AR applications is manual 3D modelling. The term “3D modelling” refers to using specialised software to construct a mathematical, coordinate-based representation of any surface of an object in three dimensions (Flavell 2010). This is most commonly achieved by manipulating and arranging or assembling vertices, edges, and polygons in a virtual 3D environment similarly to creating a mosaic. The resulting 3D model is a collection of data points in 3D space, connected by various geometric shapes and entities called a mesh. Additionally, the surface of a 3D model can be enhanced with texture mapping to add realism and detail. This approach is effective for producing low-poly models with high levels of accuracy, especially when working from a clear initial concept. Some software applications allow 3D modelling from an initial reference image; however, these approaches may not be suitable for more detailed or complex models. Some software supports working on solid models (the process often called digital sculpting, sculpt modelling, or 3D sculpting) instead of mesh (shell, boundary) polygonal modelling described above. Digital sculpting has been proposed as an alternative to pen and paper in the early stages of the design process (Alcaide-Marzal et al. 2013).

One of the main advantages of manual 3D modelling is the ability to create fine detailed and precise 3D models. However, it can also be time-consuming and limited to designers and engineers with proper technical expertise (Borrel and Fourches

2017; Livesu et al. 2017). Professions that utilise 3D modelling include product development, automotive design, industrial equipment manufacturing, architecture, engineering, entertainment, and healthcare. These fields rely on 3D modelling as a tool for creating accurate and detailed representations of objects and structures that are used for a variety of mediums including video games, movies, architecture renders, illustrations, engineering, and commercial advertising. It needs to be emphasised that all these fields could utilise AR to showcase these models. Because of the relative tediousness in manually creating 3D models, there is a strong market for these models and related content, including textures and scripts. Online platforms such as TurboSquid, CGStudio, CreativeMarket, MyMiniFactory, Sketchfab, CGTrader, ThinkVerse and Cults enable individual artists to either offer or sell their creations and monetise projects (Groenendyk, 2016). These repositories could also be used to automatically feed the AR systems on demand based on the context and task at hand.

16.3.2 Context Aware Text and Image Augmentations

The above-mentioned methods require preparing the content well in advance. Using the semantic understanding from the environment that was obtained through object detection and recognition, one can decide on what meaningful text and image augmentations could be automatically generated for a particular scenario at hand based on for example various location based services, or chat boat agents based on a large-scale language model (LLM) that can hold a conversation with the user (e.g., the aforementioned ChatGPT (Brown et al. 2020)). In this way context aware information could be ubiquitously generated for arbitrary AR scenes without the need to prepare the information in advance. Furthermore, such augmentations could be personalised to individual users, as is a common practice in most online services nowadays. Another approach would be also to enhance augmentations with synthetically generated imagery. The example given in the introduction about learning vocabulary in a foreign language would benefit from this approach. The example given mentioned the advantage of the mnemonic called the “keyword method”. Visualisation of keywords (memorable links) for arbitrary objects could be generated using 2D image synthesis. In a similar way, a context aware location-based game could be generated using text prompts and text to image generation. Various text to image methods are presented in the next subsection.

16.3.3 Text to Image Generation

Automatically creating images that semantically match a user-provided text description is a challenging problem, which proved elusive until the advent of Deep Learning, and more concretely until the advent of deep generative models. In an early work,

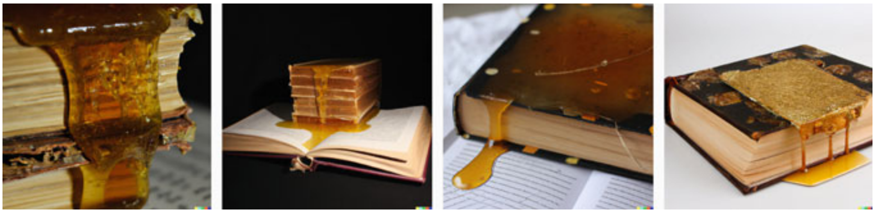


Fig. 16.6 Synthetically generated photorealistic image based on text description "a book covered with honey". The description is based on linking the Japanese word for a book—"hon"—and the familiar sounding word in English—"honey", which is a technique called the keyword method used to learn a vocabulary of a foreign language. Such synthetically generated images could be fed to the AR system presented in Fig. 16.3 of this chapter

Mansimov et al. (2015) reversed the AlignDRAW algorithm (Gregor et al. 2015) to generate images from text, but the quality of the resulting images was not good enough for real-world usage. In the last five years, different deep generative models have been devised to synthesise images from text, including Generative Adversarial Networks (GANs) (Reed et al. 2016), Variational Auto-Encoders (VAEs) (Huang et al. 2018), flow-based models (Yang et al. 2020), autoregressive models (Esser et al. 2021), and more recently diffusion models (Esser et al. 2021).

Diffusion models have recently demonstrated the ability to produce realistic images with simpler architectures compared to other models. The key idea is to generate samples from the input data distribution first by adding random noise and then progressively removing the noise. The DALL-E 2 (Ramesh et al. 2022), Imagen (Saharia et al. 2022), GLIDE (Nichol et al. 2022), and Midjourney[3] are just some of the popular text-to-image generators based on diffusion models. Finally, Stable Diffusion (Rombach et al. 2022) is a new class of latent diffusion models that provides more faithful and detailed reconstructions than previous work. In contrast to previous work, the model is trained on the learned latent space, therefore it does not rely on excessive spatial compression. The reduced complexity also provides efficient image generation from the latent space with a single network pass. The result of such a model (Ramesh et al. 2022) can be seen in Fig. 16.6, where we used a description "a book covered with honey" as a memorable link between the Japanese word for a book—"hon"—and the familiar sounding word in English—"honey".

16.3.4 Neural Rendering

There are two general types of neural rendering, namely 2D and 3D neural rendering. In 2D neural rendering, the neural network tries to convert scene parameters directly into the output image. These scene parameters are usually provided as one-dimensional vectors or as two-dimensional images (see Fig. 16.7 top). Contrary to

[3] https://www.midjourney.com..

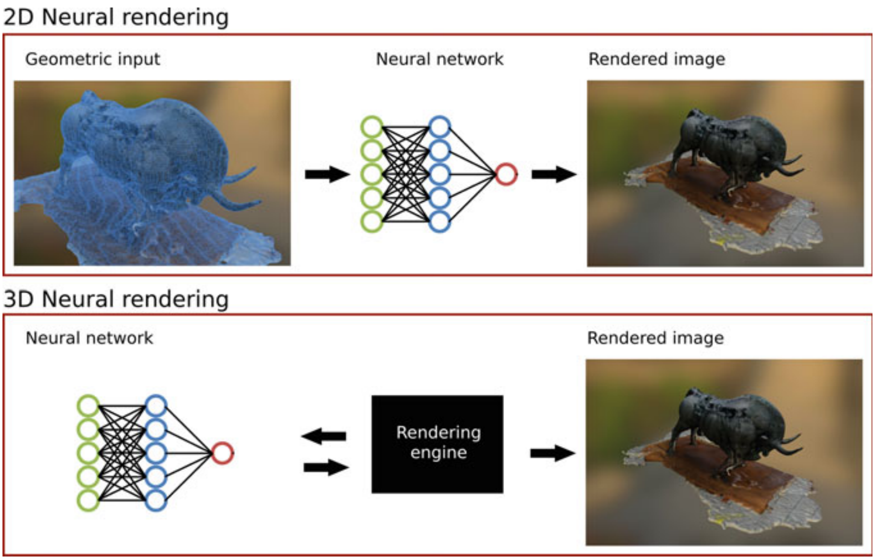


Fig. 16.7 The difference between 2 and 3D neural rendering

this direct approach, a 3D neural network attempts to find a representation of the shape or appearance of a 3D scene, which can then be rendered using a classical rendering engine such as forward rendering or ray casting. This rendering engine is not learnt but is defined analytically (see Fig. 16.7 bottom). One of such 3D neural rendering methods is NeRF, which learns physically meaningful colour and density values in 3D space (the output) whilst it uses camera pose and 2D image data as an input to the neural network. As indicated in Fig. 16.8, learning is based on minimising the rendering loss. These representations can then be rendered through ray casting and volume integration generating novel views of the scene (Tewari et al. 2022). This approach can be used for view synthesis of static and dynamic scenes as well as the generation of new scenes. We discuss each in subsections hereafter.

16.3.4.1 View Synthesis of Static Scenes

The overarching goal of view synthesis is to take a set of scene images and render a given scene from new camera positions. Besides camera movement, it also involves editing the scene, such as, for example, moving objects around or changing the lighting conditions. However, the important part here is that input images come from a static scene with no or minimal movement as well as constant lighting conditions. There are two general approaches that have been followed in the literature: view synthesis from 3D voxel grids (Flynn et al. 2019; Mildenhall et al. 2019; Zhou et al. 2018) and view synthesis from neural network representations (Barron et al. 2022; Mildenhall et al. 2022; Müller et al. 2022). Voxel grid representations were used first,

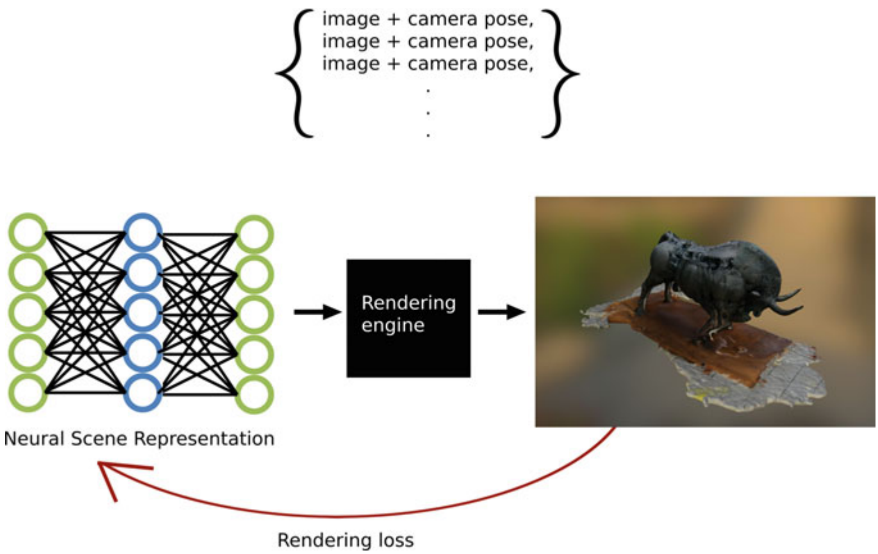


Fig. 16.8 Neural radiance field (NeRF) scene representation and training

however, they consume a lot of memory, which greatly limits the resolution of such generated imagery. Contrary, neural network representations require less memory, which have seen significant performance and quality gains in recent times.

The performance of view synthesis is commonly evaluated based on rendering realism, consistency of views as the camera moves from one point of view to the other, and finally the amount of data and time that it takes to generate representations and render these views. In recent times, there have been significant advantages in all of these areas showing great promise for neural rendering.

16.3.4.2 View Synthesis of Dynamic Scenes

Beside synthesising views for static scenes, it is also possible to synthesise views of dynamic scenes. In such scenarios, one needs to find a way to generate dynamic neural scene representation, which can be done implicitly by conditioning the neural network to deformation states, or explicitly via a separate deformation representation. This can, amongst others, give us control over deformable states and allow us for example to synthesise novel views of expressions on human faces (Gafni et al. 2021). Further still, this enables us to generate head avatars (Cao et al. 2022; Zielonka et al. 2022) that can be controlled in real time. Cao et al. achieve this by building an avatar model that can be modified to novel identities using only a phone scan of a neutral face expression (Cao et al. 2022). With this technique we can create controllable human-like head avatars that go beyond the uncanny valley of human resemblance.

16.3.4.3 Generalisation of Scene Representation

One approach of synthesising views is to fit the neural scene representation to a single scene. However, this generally requires a large number of views. In this subsection, we discuss approaches that can generalise across multiple scenes and thus enable synthetic image generation from a limited set of views. There are two lines of work in achieving this: local and global conditioning (Tewari et al. 2022). One such example is the LOLNeRF system in which authors train the model using solely a single-view image of a person that can then be rendered from multiple views (Rebain et al. 2022). In other worlds, a conversion from a single image of a person's face to a 3D viewing experience is generated. Another interesting example of generalising on humans is the HumanNeRF system that works on a given monocular video of a human performing complex body motions. The system is able to render a person in the video from an arbitrary point of view at any given time within the video (Weng et al. 2022).

Another interesting approach generalising NeRF is to bind it with natural language to processing through which 3D objects are generated (Jain et al. 2022; Mohammad Khalid et al. 2022). Yet another example of generalisation of scene representations is DreamFusion in which 2D diffusion models are used to generate images based on textual descriptions that are grouped with NeRF in order to achieve rendering from multiple camera views (Poole et al. 2022). A similar approach to DreamFusion is Magic3D, which achieves significantly higher resolution at faster runtimes (Lin et al. 2022). In the case of GAUDI, the generation focuses on large scene generation. It uses text or image conditioning, which generates scenes and camera paths (Bautista et al. 2022). Another approach of generalisation is to use multimodal conditions for 3D shape completion, reconstruction or generation as was accomplished with SDFusion (Cheng et al. 2022).

16.4 Discussion and Conclusion

By envisioning a combination of novel AI services with AR systems, we provide an example of use and interaction between the AI powered AR system and the user. We use vocabulary learning through AR HMD as an example (see Fig. 16.1). To recap, the system takes advantage of the keyword method to reinforce the memorisation of foreign words. The user needs to find a familiarly sounding word (e.g. in their own language) to the word being learned, and come up with the association (a memorable link) between the object (of the word being learnt) and familiarly sounding word/phrase.

The user is wearing the HMD glasses, sees an object and points to it. The following interaction might unfold like this:

1. The system detects the object, which is a postcard.
2. The system performs semantic and instance segmentation of the postcard.

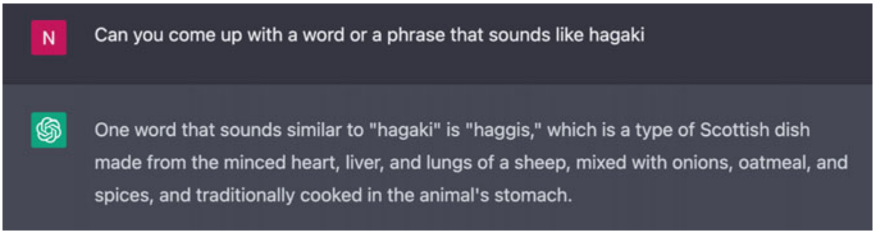


Fig. 16.9 An example of the chat with ChatGPT[4] about finding the keyword used in the keyword method to reinforce learning vocabulary of a foreign language

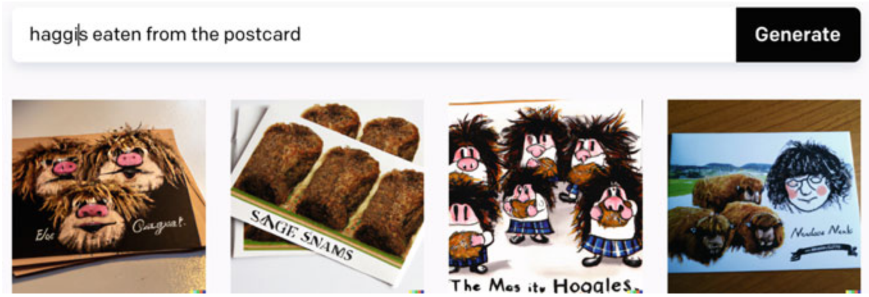


Fig. 16.10 An example of images returned by the DALL-E 2[5] system based on the text provided by the user

3. The system queries a translation service to obtain the word for a postcard in a foreign (Japanese) language ("hagaki").
4. The system uses speech synthesis to get the pronunciation of the word.
5. The system provides the word and the pronunciation to be played to the user.
6. The user plays the pronunciation but cannot find a familiarly sounding word/phrase to generate the keyword and the memorable link.
7. The user queries an AI language model chatbot (e.g. ChatGPT) for a word or a phrase that sounds familiar to the Japanese word "hagaki" (see Fig. 16.9).
8. The system returns the word "haggis".
9. The user provides a memorable link to the text-to-image service "haggis eaten from the postcard".
10. The system returns a set of synthetically generated images (see Fig. 16.10).
11. The user selects one that suits their memorable link
12. The system places it besides the word and pronunciation.

The instance of the content is saved for future use. Whenever the user points to a postcard, the content is shown by it.

[4] OpenAI's ChatGPT https://chat.openai.com/.
[5] OpenAI's DALL-E 2 https://openai.com/dall-e-2/.

However, using these novel AI techniques comes with several caveats if they are to be used in AR systems. The availability of AI language models offered to the public in the form of chatbots (such as ChatGPT mentioned in this chapter), is an exciting advancement of the field but it also raises some concerns. First, one is the ownership of the output produced by such systems since the AI language models are trained on publicly available text (often collected without consent) and may reproduce this information verbatim. However, there is an even bigger concern, which is the factuality of the text produced. ChatGPT and similar autocomplete systems are designed to predict the next word in a sentence based on the words that come before it. These tools do not have a set database of facts embedded and rely on their ability to generate believable statements. This can lead to the presentation of false information as true. Nevertheless, in the scenario presented above, it is up to the user to decide if the keyword is the one they might use to generate the memorable link. Moreover, these systems might not produce the desired results in the first try if at all. But the models will likely get better with time.

Despite significant advances of neural rendering in recent years, these methods are still faced with significant challenges for adoption within AR systems. One of the key challenges is the speed of neural rendering. Commonly available graphic cards are extremely fast at rendering polygon meshes, which makes it possible to create high resolution synthetic images with high refresh rates and very low latency. Neural rendering approaches have a hard time achieving comparable real time performance, which is essential in AR systems, especially as real time performance is a must. Another challenge is generalisation of neural scene representations. For example, in neural representations of the scene it is usually impossible to adjust lighting or material properties. This is problematic because realistic rendering of 3D graphics in real world environments does not solely rely on the correct geometry and registration, but also on correctly adjusting to the lighting conditions within the scenes. However, the models will likely get better with time.

Another important element is scene manipulation, such as object movement and deformation. These are yet again very important basic functionalities of the 3D scenes in AR where augmented objects are animated in various ways in order to achieve the desired visual effects. Further still, current work in neural rendering mainly focuses on single objects or scenes with relatively low complexity. This is yet again limiting the applicability of neural rendering to AR.

Despite all the above-mentioned shortcomings, the presented AI models offer a unique new opportunity to increase generalisability of AR services, which could drastically increase the possibilities of future AR systems and are therefore very important for future advances in AR.

Acknowledgements The authors also acknowledge the European Commission for funding the InnoRenew CoE project (Grant Agreement 739574) under the Horizon2020 Widespread-Teaming program and the Republic of Slovenia (Investment funding of the Republic of Slovenia and the European Union of the European Regional Development Fund). The research was also supported by Slovenian research agency ARRS (P1-0383 Complex networks, P5-0433 Digital restructuring of deficient professions for society 5.0 (Industry 4.0), J1-9186, J1-1715, J5-1796 and J1-1692).

References

Alcaide-Marzal J, Diego-Más JA, Asensio-Cuesta S, Piqueras-Fiszman B (2013) An exploratory study on the use of digital sculpting in conceptual product design. Des Stud 34(2):264–284. https://doi.org/10.1016/j.destud.2012.09.001

Barron JT, Mildenhall B, Verbin D, Srinivasan PP and Hedman P (2022) Mip-NeRF 360: unbounded anti-aliased neural radiance fields. (arXiv:2111.12077). arXiv. http://arxiv.org/abs/2111.12077

Bautista MA, Guo P, Abnar S, Talbott W, Toshev A, Chen Z, Dinh L, Zhai S, Goh H, Ulbricht D, Dehghan A, Susskind J (2022) GAUDI: a neural architect for immersive 3D scene generation. (arXiv:2207.13751). arXiv. http://arxiv.org/abs/2207.13751

Bemis SP, Micklethwaite S, Turner D, James MR, Akciz S, Thiele ST, Bangash HA (2014) Ground-based and UAV-Based photogrammetry: a multi-scale, high-resolution mapping tool for structural geology and paleoseismology. J Struct Geol 69:163–178. https://doi.org/10.1016/j.jsg.2014.10.007

Bi H, Zheng W, Ren Z, Zeng J, Yu J (2017) Using an unmanned aerial vehicle for topography mapping of the fault zone based on structure from motion photogrammetry. Int J Remote Sens 38(8–10):2495–2510. https://doi.org/10.1080/01431161.2016.1249308

Borrel A, Fourches D (2017) RealityConvert: a tool for preparing 3D models of biochemical structures for augmented and virtual reality. Bioinformatics 33(23):3816–3818. https://doi.org/10.1093/bioinformatics/btx485

Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler DM, Wu J, Winter C (2020) Language models are few-shot learners. In: Proceedings of the 34th international conference on neural information processing systems

Cao C, Simon T, Kim JK, Schwartz G, Zollhoefer M, Saito S-S, Lombardi S, Wei SE, Belko D, Yu SI, Sheikh Y, Saragih J (2022) Authentic volumetric avatars from a phone scan. ACM Trans Graphics 41(4):1–19. https://doi.org/10.1145/3528223.3530143

Cheng YC, Lee HY, Tulyakov, S, Schwing A, Gui L (2022) SDFusion: multimodal 3D shape completion, reconstruction, and generation. (arXiv:2212.04493). arXiv. http://arxiv.org/abs/2212.04493

Collins T, Woolley SI, Gehlken E, Ch'ng E (2019) Automated low-cost photogrammetric acquisition of 3D models from small form-factor artefacts. Electronics 8(12):1441. https://doi.org/10.3390/electronics8121441

Cui Y, Schuon S, Chan D, Thrun S, Theobalt C (2010) 3D shape scanning with a time-of-flight camera. IEEE Comput Soc Conf Comput vis Patt Recog 2010:1173–1180. https://doi.org/10.1109/CVPR.2010.5540082

Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1, pp 886–893. https://doi.org/10.1109/CVPR.2005.177

Enesi I, Kuqi A, Zanaj E (2022) Quality of 3D reconstruction based on photogrammetry for small objects, a case study. IOP Conf Ser: Mater Sci Eng 1254(1):012039. https://doi.org/10.1088/1757-899X/1254/1/012039

Esser P, Rombach R, Blattmann A, Ommer B (2021) ImageBART: Bidirectional context with multinomial diffusion for autoregressive image synthesis. In: Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Vaughan JW (eds), Advances in neural information processing systems, vol 34. Curran Associates, Inc. pp 3518–3532. https://proceedings.neurips.cc/paper/2021/file/1cdf14d1e3699d61d237cf76ce1c2dca-Paper.pdf

Flavell L (2010) Beginning Blender: open source 3D modeling, animation, and game design. Apress, Distributed to the book trade worldwide by Springer Science+Business Media

Flynn J, Broxton M, Debevec P, DuVall M, Fyffe G, Overbeck R, Snavely N, Tucker R (2019) DeepView: view synthesis with learned gradient descent. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2019:2362–2371. https://doi.org/10.1109/CVPR.2019.00247

Fried P, Woodward J, Brown D, Harvell D, Hanken J (2020) 3D scanning of antique glass by combining photography and computed tomography. Digit Appl Archaeol Cult Heritage 18:e00147. https://doi.org/10.1016/j.daach.2020.e00147

Gafni G, Thies J, Zollhofer M, Niesner M (2021) Dynamic neural radiance fields for monocular 4D facial avatar reconstruction. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2021:8645–8654. https://doi.org/10.1109/CVPR46437.2021.00854

Gregor K, Danihelka I, Graves A, Rezende DJ, Wierstra D (2015) DRAW: A recurrent neural network for image generation. In: Proceedings of the 32nd international conference on international conference on machine learning, vol 37, pp 1462–1471

Groenendyk M (2016) Cataloging the 3D web: the availability of educational 3D models on the internet. Library Hi Tech 34(2):239–258. https://doi.org/10.1108/LHT-09-2015-0088

Hall T, Schnädelbach H, Flintham M, Ciolfi L, Bannon L, Fraser M, Benford S, Bowers J, Greenhalgh C, Hellström SO, Izadi S (2001) The visitor as virtual archaeologist: explorations in mixed reality technology to enhance educational and social interaction in the museum. In: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage—VAST '01, vol 91. https://doi.org/10.1145/584993.585008

Huang H, Li Z, He R, Sun Z, Tan T (2018) IntroVAE: introspective variational autoencoders for photographic image synthesis. In: Proceedings of the 32nd international conference on neural information processing systems, pp 52–63

Ibrahim A, Huynh B, Downey J, Hollerer T, Chun D, O'donovan J (2018) ARbis pictus: a study of vocabulary learning with augmented reality. IEEE Trans Visual Comput Graphics 24(11):2867–2874. https://doi.org/10.1109/TVCG.2018.2868568

Izadi S, Kim D, Hilliges O, Molyneaux D, Newcombe R, Kohli P, Shotton J, Hodges S, Freeman D, Davison, Fitzgibbon A (2011) KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In: Proceedings of the 24th annual ACM symposium on user interface software and technology, pp 559–568. https://doi.org/10.1145/2047196.2047270

Jain A, Mildenhall B, Barron JT, Abbeel P, Poole B (2022) Zero-shot text-guided object generation with dream Fields. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2022:857–866. https://doi.org/10.1109/CVPR52688.2022.00094

Lin CH, Gao J, Tang L, Takikawa T, Zeng X, Huang X, Kreis K, Fidler S, Liu MY, Lin TY (2022) Magic3D: high-resolution text-to-3D content creation. (arXiv:2211.10440). arXiv. http://arxiv.org/abs/2211.10440

Livesu M, Ellero S, Martínez J, Lefebvre S, Attene M (2017) From 3D models to 3D prints: an overview of the processing pipeline. Comput Graphics Forum 36(2):537–564. https://doi.org/10.1111/cgf.13147

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vision 60(2):91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Mansimov E, Parisotto E, Ba JL, Salakhutdinov R (2015) Generating images from captions with attention. https://doi.org/10.48550/ARXIV.1511.02793

Mikhail EM, Bethel JS, McGlone JC (2001) Introduction to modern photogrammetry. Wiley

Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorth R, Ng R (2022) NeRF: Representing scenes as neural radiance fields for view synthesis. Commun ACM 65(1):99–106. https://doi.org/10.1145/3503250

Mildenhall B, Srinivasan PP, Ortiz-Cayon R, Kalantari NK, Ramamoorthi R, Ng R, Kar A (2019) Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. ACM Trans Graphics 38(4):1–14. https://doi.org/10.1145/3306346.3322980

Mohammad Khalid N, Xie N, Belilovsky E, Popa T (2022) CLIP-Mesh: generating textured meshes from text using pretrained image-text models. In: SIGGRAPH Asia 2022 conference papers, pp 1–8. https://doi.org/10.1145/3550469.3555392

Müller T, Evans A, Schied C, Keller A (2022) Instant neural graphics primitives with a multiresolution hash encoding. https://doi.org/10.48550/ARXIV.2201.05989

Nichol A, Dhariwal P, Ramesh A, Shyam P, Mishkin P, McGrew B, Sutskever I, Chen M (2022) GLIDE: towards photorealistic image generation and editing with text-guided diffusion models. (arXiv:2112.10741). arXiv. http://arxiv.org/abs/2112.10741

Pepe M, Costantino D (2020) Techniques, tools, platforms and algorithms in close range photogrammetry in building 3D model and 2D representation of objects and complex architectures. Comput-Aid Design Appl. 18(1):42–65. https://doi.org/10.14733/cadaps.2021.42-65

Poole B, Jain A, Barron JT, Mildenhall B (2022) DreamFusion: Text-to-3D using 2D diffusion (arXiv:2209.14988). arXiv. http://arxiv.org/abs/2209.14988

Qi W, Li F, Zhenzhong L (2010) Review on camera calibration. Chin Control Decis Conf 2010:3354–3358. https://doi.org/10.1109/CCDC.2010.5498574

Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M (2022) Hierarchical text-conditional image generation with CLIP latents. (arXiv:2204.06125). arXiv. http://arxiv.org/abs/2204.06125

Rebain D, Matthews M, Yi KM, Lagun D, Tagliasacchi A (2022) LOLNeRF: Learn from one look. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2022:1548–1557. https://doi.org/10.1109/CVPR52688.2022.00161

Reed S, Akata Z, Yan X, Logeswaran L, Schiele B, Lee H (2016) Generative adversarial text to image synthesis. In: Proceedings of the 33rd international conference on international conference on machine learning, vol 48, pp 1060–1069

Remondino F (2011) Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sens 3(6):1104–1138. https://doi.org/10.3390/rs3061104

Ren J, Wang Y (2022) Overview of object detection algorithms using convolutional neural networks. J Comput Commun (10):115–132. https://doi.org/10.4236/jcc.2022.101006

Rocchini C, Cignoni P, Montani C, Pingi P, Scopigno R (2001) A low cost 3D scanner based on structured light. Comput Graphics Forum 20(3):299–308. https://doi.org/10.1111/1467-8659.00522

Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B (2022) High-resolution image synthesis with latent diffusion models. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2022:10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042

Saharia C, Chan W, Saxena S, Li L, Whang J, Denton E, Ghasemipour SKS, Ayan BK, Mahdavi SS, Lopes R G, Salimans T, Ho J, Fleet DJ, Norouzi M (2022) Photorealistic text-to-image diffusion models with deep language understanding. (arXiv:2205.11487). arXiv. http://arxiv.org/abs/2205.11487

Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B (2020) Google Earth Engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J Photogramm Remote Sens 164:152–170. https://doi.org/10.1016/j.isprsjprs.2020.04.001

Tewri A, Thies J, Mildenhall B, Srinivasan P, Tretschk E, Yifan W, Lassner C, Sitzmann V, Martin-Brualla R, Lombardi S, Simon T, Theobalt C, Nießner M, Barron JT, Wetzstein G, Zollhöfer M, Golyanik V (2022) Advances in neural rendering. Comput Graphics Forum 41(2):703–735. https://doi.org/10.1111/cgf.14507

Weerasinghe M, Biener V, Grubert J, Quigley A, Toniolo A, Pucihar KC, Kljun M (2022) VocabulARy: learning vocabulary in AR supported by keyword visualisations. IEEE Trans Visual Comput Graphics 28(11):3748–3758. https://doi.org/10.1109/TVCG.2022.3203116

Weng CY, Curless B, Srinivasan PP, Barron JT, Kemelmacher-Shlizerman I (2022) HumanNeRF: free-viewpoint rendering of moving people from monocular video. IEEE/CVF Conf Comput vis Patt Recog (CVPR) 2022:16189–16199. https://doi.org/10.1109/CVPR52688.2022.01573

Wulder MA, White JC, Nelson RF, Næsset E, Ørka HO, Coops NC, Hilker T, Bater CW, Gobakken T (2012) Lidar sampling for large-area forest characterization: a review. Remote Sens Environ 121:196–209. https://doi.org/10.1016/j.rse.2012.02.001

Yang K, Goldman S, Jin W, Lu A, Barzilay R, Jaakkola T, Uhler C (2020) Improved conditional flow models for molecule to image synthesis. https://doi.org/10.48550/ARXIV.2006.08532

Zhou T, Tucker R, Flynn J, Fyffe G, Snavely N (2018) Stereo magnification: Learning view synthesis using multiplane images. ACM Trans Graphics 37(4):1–12. https://doi.org/10.1145/3197517.3201323

Zielonka W, Bolkart T, Thies J (2022) Instant volumetric head Avatars. (arXiv:2211.12499). arXiv. http://arxiv.org/abs/2211.12499

Chapter 17
Smart Extended Reality in the Metaverse-Tailing: The Rise of New Retail Landscape

Federica Caboni and Lucia Pizzichini

Abstract This chapter aims to investigate the rise of Metaverse by considering the merge of interactive technologies (such as augmented and virtual reality and artificial intelligence) into a new retail landscape. The research flow of this chapter is guided by three theories able to identify and clarify how a smart Extended Reality in the Metaverse-Tailing can be considered as the rise of new retail environment. First, the affordance theory of technology will help to understand the possibility of obtaining value from specific technology, deriving from the Metaverse world. Secondly, the lens of regulatory engagement theory is helpful to understand the positive engagement of people during their experience in a specific object or environment thanks to the exploitation of augmented and virtual reality in a Metaverse world. In line with the regulatory engagement theory, this chapter aims to underline the possible way of involvement in the Metaverse during a shopping journey and exploiting potentialities deriving from Augmented Reality, Virtual Reality and Artificial Intelligence, such as interaction, immersion, inspiration and satisfaction. Finally, the self-determination theory identifies the possible intrinsic or extrinsic motivation that leads people to experience retailing in the metaverse-tailing as a fusion of several advanced technologies.

17.1 Introduction

In the last few years, the emergent technologies (Grewal et al. 2021, 2020a) that are modifying retail settings are imposing profound reflections on how it is possible to

F. Caboni (✉)
Department of Management, Alma Mater Studiorum, University of Bologna, Bologna, Emilia-Romagna, Italy
e-mail: f.caboni@unibo.it

L. Pizzichini
Department of Management, School of Economics "G. Fuà", Polytechnic University of Marche, Marche, Italy
e-mail: l.pizichini@staff.univpm.it

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_17

adapt to these changes. Particularly, on one side, retailers and practitioners need to re-think their strategies (Grewal et al. 2021, 2020b) to adapt to these technologies and exploit them in the best possible way to face the rise of new retail settings such as the Metaverse. On the other side, consumers are constantly involved in new shopping experiences considering several external factors. For example, the rapid spread of COVID-19 modified the way of shopping and accelerated the adoption of technologies such as augmented and virtual reality (Caboni and Pizzichini 2022) and artificial intelligence in general. The development of always new desires and needs pushes people to explore interactive technologies' potentiality, and consequently, the shopping experience continues in evolution, adapting to the continual reshaping of retail. In this chapter, the approach to the Metaverse considers several interactive technologies such as the Internet, social media, mobile technologies, augmented reality (AR), and virtual reality (VR). In this regard, people can have practical access to information and consumption channels (Shankar et al. 2021) anywhere any time, increasing their endless shopping experience. Finally, thanks to the exploitation of interactive technologies, it is simple to imagine how the shopping experience assumes a new dimension where people can engage in virtually seamless connections with retailers, manufacturers, consumers, and influencers (Dolbec and Fischer 2015; Grewal et al. 2017). People in the realm of emergent, intelligent and interactive technologies (Grewal et al. 2021, 2020a) have at their disposal the possibility to share not only shopping information with others but also decisions and evaluations on products and services, creating huge shopping community both physically and virtually. In this scenario appears fundamental to consider the new and modified ways of shopping often accelerated by the adoption of technologies such as augmented and virtual reality and artificial intelligence in general. In this venue people can explore interactive technologies' potentiality and adapting to the continual reshaping of retail. In this dynamic context, this chapter approaches the rise of Metaverse by considering the fusion of interactive technologies such as augmented and virtual reality and artificial intelligence into a new retail landscape. The research flow of this chapter is guided by three theories (Fig. 17.1) able to identify and clarify how a smart Extended Reality in the Metaverse-Tailing can be considered as the rise of new retail landscape. In particular, this chapter aims to answer to the following research questions:

RQ1: How it is possible to obtain value from the Metaverse?
RQ2: How to stimulate a positive engagement in the Metaverse?
RQ3: What intrinsic or extrinsic motivation, may lead people in the Metaverse experience?

To answer to the above-mentioned questions, this chapter will consider the affordance theory of technology (Gaver 1991), the regulatory engagement theory (Higgins and Scholer 2009), and finally the self-determination theory (Deci and Ryan 1980; 1985a; 1985b; 1987; 2000, 2012). Specifically, the affordance theory of technology (Gaver 1991) will help to understand the possibility of obtaining value from specific

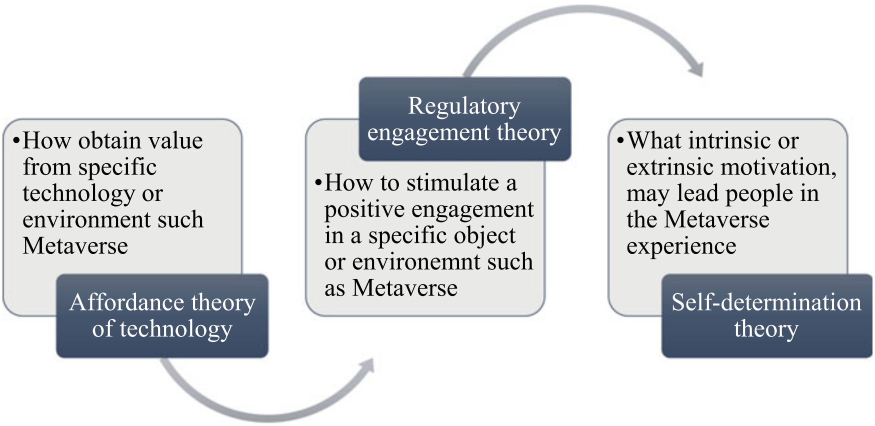


Fig. 17.1 Research flow

technology, deriving from the Metaverse world. Then, the lens of regulatory engagement theory (Higgins and Scholer 2009) is helpful to understand the positive engagement of people during their experience in a specific object or environment thanks to the exploitation of augmented and virtual reality and artificial intelligence, in a Metaverse world. In line with the regulatory engagement theory (Higgins and Scholer 2009), this chapter aims to underline the possible way of involvement in the Metaverse during a shopping journey and exploiting potentialities deriving from the fusion of augmented, virtual reality and artificial intelligence, such as interaction, immersion, inspiration and satisfaction. Finally, the self-determination theory (Deci and Ryan 1980; 1985a; 1985b; 1987; 2000, 2012) identifies the possible intrinsic or extrinsic motivation that led people to experience retailing in the Metaverse-tailing as a fusion of several advanced technologies.

17.2 Theoretical Background

To approach the rise of a new retail landscape such as the Metaverse-tailing, consider the merge of interactive technologies such as augmented and virtual reality and artificial intelligence into a new retail landscape. Notably, the section below will be presented a general description of the Metaverse by following the three theories used in this chapter to understand more deeply the smart extended reality in the Metaverse-tailing.

17.2.1 Smart Extended Reality in the Metaverse-Tailing

Identifying a smart extended reality appears particularly innovative and interesting, considering the rapid and new development of what scholars contemplating as the Metaverse. The term Metaverse was coined 30 years ago when Neal Stephenson, in his fiction novel called "Snow Crash" used the word Metaverse (Papagiannidis et al. 2008; Wright et al. 2008; Marmaridis and Griffith 2009; Sourin 2017; Murray 2020; Key et al. 2021; Park and Kim 2022), and several definitions in the course of time has been developed (see Table 17.1). More specifically, this term was used to identify a place based on one principal feature: the co-presence of real and virtual elements (Joeng 2022). Basically, from the analysis of this term derive two subordinate words: one prefix as "Meta" with the meaning of beyond and a suffix "verse" that refers to the universe (Marmaridis and Griffith 2009; Lee et al. 2011; Dionisio 2013; Fang et al. 2021; Jeon 2021; Kye et al. 2021; Jeong 2022). So literally, a meta world could be considered in this chapter as a smart world because it is composed of smart elements such as interactive and intelligent technologies and innovative services.

Specifically, the academic literature until this moment categorized the Metaverse as a place composed of interactive technologies such as Augmented reality, Mirror worlds, and Virtual worlds (Park and Kim 2022; Nevelsteen 2018; Lee et al. 2011). The value of this world rises drastically with the younger generation, such as generation Z that has a considerable attitude toward using intelligent technologies (Park and Kim 2022). At this moment, the academic literature with a specific reference to business and management still needs to be developed, and there is a paucity of studies able to identify how this meta-world is developing. However, the 2021 can be considered a milestone because Facebook (now called Meta) started to invest several amounts of money in developing its Metaverse strategy (Binson 2021; Rauschnabel et al. 2022; Rospigliosi 2022; Wiederhold 2022). This example allows us to glimpse how, in the coming years, the Metaverse will constitute a great source of economic, commercial and social value (Papagiannidis 2008; Daz et al. 2020; Park and Kim 2022) for different categories of people and companies that will be present in the Metaverse. In fact, after Facebook's investment on the Metaverse (Binson 2021; Rauschnabel et al. 2022; Rospigliosi 2022; Wiederhold 2022), people started to express their interest in this murky world. Moreover, it might be desirable to think that shortly a smart extended reality will be the first-place where people will go to experience immersive experiences (Papagiannidis 2008). Notably, thanks to the support of intelligent and smart technologies, people can socialize and interact (Wright et al. 2008) in a parallel world such as the Metaverse. Since the rise of the words Metaverse, scholars have developed several attempts to define and analyse the different features of this technology. Scholars refer to this world as a combination of virtual objects in a natural environment (Lee et al. 2011; Davis et al. 2009; Gadalla 2013; Jeong 2022). And as a virtual world where people can interact in real-time with others thanks to the support of an avatar (Leenes 2008; Davis et al. 2009; Owens et al. 2011; Gadalla 2013; Daz et al. 2020; Laviola et al. 2022; Wiederhold 2022). All in all, the main characteristics of the Metaverse can be identified in the following points (summarised in Fig. 17.2):

Table 17.1 Principal Metaverse definitions

Authors and year	Definitions
Stephenson 1992	A word where humans as avatars can interact with others in a 3 dimensions space that reflect a physical world
Jaynes et al. 2003	An immersive environment with digital media network able to remove the barriers of time and space
Rymaszewski et al. 2007	An environment where people are able to create their personality, visit different places, explore expansive buildings and shop
Collins 2008	An interactive network with continuous, immersive 3D virtual environment accessible
Wright et al. 2008	Extensive 3D virtual world able to support people for social interaction
Schlemmer et al. 2009	Extension of physical world within the virtual internet
Schaf et al. 2009	A world of enhancing the feeling of being in a classroom
Messinger et al. 2009	A virtual world where thousands of people can interact simultaneously within the same 3D environment
Cunningham 2010	A compound word of meta and universe where the real and virtual world are mixed
Owens et al. 2011	An immersive 3D virtual world where people interact each-others by using real worlds metaphors but without physical limitations
Toneis 2011	A world that reconstructs the meaning of the living world with experience
Guo et al. 2011	A computer simulation that allow avatars to interconnect and communicate in relatively life-like environments
Kim et al. 2013	A collective online space created by combining physical reality enhanced by a 3Dvirtual world and a physical permanent virtual space
Luse et al. 2013	A virtual world that allow people to live their virtual life online
Dionisio et al. 2013	An integrated network of 3D virtual worlds in an independent virtual world or an attractive realm for human sociocultural interaction
Papagiannidis et al. 2014	A place where users are able to create content and object they want
Preda et al.	Collective online shared environment
Ko and Jang 2014	An online virtual community that allow people to interact each-others through avatars
Dascalu et al. 2014	New environment where physical and digital objects co-exist
Amorim et al. 2014	An immersive environment that can simulate real world features
Chen 2016	AN immersive environment that reflect the real world co-created by users
Choi and Kim 2017	A space created by the fusion of virtual and augmented reality
Huggett 2020	A world where virtual worlds combine immersive VR with physical actors, objects, interfaces and networks in a future form of the Internet

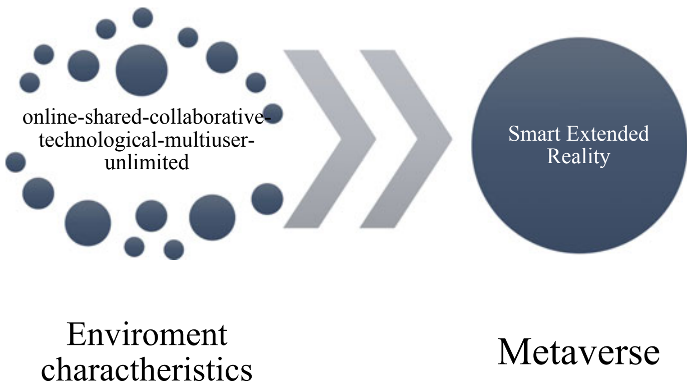


Fig. 17.2 Smart extended reality

(1) A collaborative environment (Kumar et al. 2008; Wright et al. 2008; Gadalla 2013) because people collaborate to achieve several different goals such as economic, social and leisure (Kumar et al. 2008).
(2) An online place because people are immersed in an online environment based on three dimensions (Papagiannidis 2008; Marmaridis and Griffith 2009; Owens et al. 2011; Gadalla 2013; Kawaguchi et al. 2020; Binson 2021).
(3) A shared world (Kim 2021) because people have the possibility to share their activities, opinion, and information (Binson 2021; Davis et al. 2009; Papagiannidis 2008) and the shopping becomes a networked experience (Pantano and Gandini 2018).
(4) An augmented and technological place (Huggett, 2020) because people thanks to the support of augmented reality have the possibility to augment their experience (Caboni and Hagberg 2019; Caboni and Pizzichini 2022), and also socialize and interact (Wright et al. 2008) thanks to the support of virtual elements, technologies and the Internet (Han et al. 2010a, b).
(5) A multiuser environment (Daz et al. 2020) because people can use the same technologies or conduct the same activities at the same time, as an extension of their real life (Kuam et al. 2008).
(6) An unlimited world because the Metaverse could be the reproduction a physical world but without any physical limitations (Marmaridis and Griffith 2009; Papagiannidis 2008; Daz et al. 2020; Kawaguchi et al. 2020; Leenes 2008).

17.3 Affordance Theory of Technology

The term affordance introduced by Gibson in 1977, refers to the invitation to use an object that through its physical quality suggests people how to use it. Specifically, Gibson (1977; 1979) explained how animals, without resorting to any kind of reasoning, were able to grasp the intrinsic meaning of an object using simply sensory perception. Hence, the concept of affordance refers to the perception, originated by the senses, that a user has of the relationship that is established between him and an object within an environment. After Gibson (1977; 1979), Norman (1988; 1999), highlighted another point of view by studying the perceptions that an object transmits to an actor and from which it is possible to define a product architecture that is intuitive and usable. According to these two principal concepts related the affordance and extending those in the Metaverse environment, the affordances deriving from the Metaverse are related to a user relationship with the Metaverse environment obtaining specific value from the exploitation of technologies in this world. In this context, the application of the concept of affordance to the technology comes from Gaver (1991) that stated that affordances are properties of the world that pay off possible some action to an organism to act in specific ways. Hence, the affordance theory of technology (Gaver 1991) appears useful to understand how it is possible to obtain value from the use of technology and more specifically from the Metaverse environment considered as a melting pot of several technologies such as augmented reality, virtual reality and internet of thing, or artificial intelligence more in general. In particular, as expressed by Gaver (1991) the affordance theory of technology refers to using a specific technology under specific environmental circumstances to address particular goals.

In the Metaverse environment, it is possible to identify specific conditions and circumstances from which people can obtain value (Gibson 2014) and satisfying their needs and desires. As this world is online, shared, unlimited, collaborative, technological and multiuser environment, people can satisfy their needs for an immersive experience and obtain value (Gibson 2014) from the interactive technologies (Pantano 2016; Pantano and Gandini 2017). In particular a world described as a smart extended reality could be characterized by several stimuli deriving from the use of immersive and interactive technologies (Pantano 2016) that are able to influence the consumers behaviour's (Pantano and Gandini 2017). In this regard the affordance theory of technology (Gaver 1991) applied to the Metaverse make it possible to consider this place as a plethora of technologies (Augmented, Virtual, Artificial) with different affordances that permit people to make possible their actions as in a normal life. In fact, according to the affordance theory, the focus to understand how to get a value from the Metaverse should be not properly on the technologies but on the fundamental interactions between users, context, environment and technologies (Fig. 17.3). In accord to the affordance theory of technology (Gaver 1991) the value from the Metaverse derives from the connection of the elements composing the Metaverse such as users (real and avatars), the context (technological and digital) where people take place their actions, and the environment (shared, collaborative, multiuser, unlimited) where they conduct their experience.

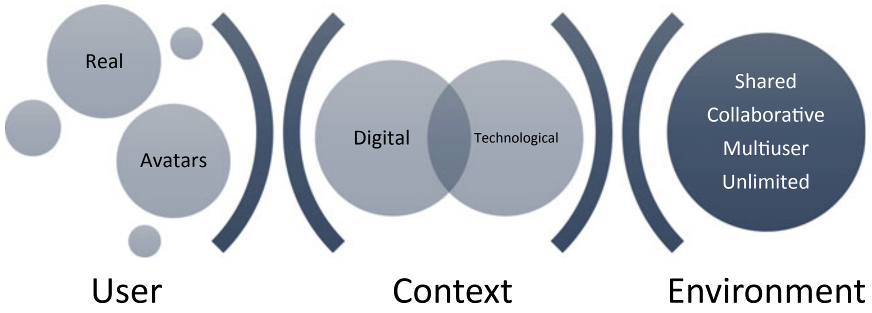


Fig. 17.3 The affordance theory in the Metaverse

17.4 Regulatory Engagement Theory

The engagement is another element that people (firms and practitioners more in general) involved in the creation and exploitation of a Metaverse need to take into consideration to create as much as possible an immerse experience (Pantano 2016; Pantano and Gandini 2017) for users. Particularly, considering the regulatory engagement theory (RET) the psychological state of a person related to the engagement (Higgins 2006; Higgins and Scholer 2009), is referred the attention stimulated by an attractive or repulsive motivational factors. In fact, a positive engagement (such in the Metaverse environment) is able to induces and attractive experience and on the contrary if the experience has a negative value people are prompt to reject the experience (Kuvykaite and Tarute 2015). More in depth, this theory appears useful to discern the motivational factors (strong or weak) that lead in their experience and the directions they take, towards or away (Higgins and Scholer 2009). In this way it is possible to predict the consumers behaviour in a new environment and understand the process followed by user in co creation of value (Scholer and Higgins 2009) in the Metaverse. Accordingly, to the regulatory engagement theory (Higgins 2006; Higgins and Scholer 2009) a positive engagement of an object, such as the Metaverse, can produce an engaging experience (Arghashi and Yuksel 2022). A strong engagement (Higgins and Scholer 2009), deriving from the immersion in the Metaverse world by using different kind of smart technologies (Augmented, virtual, artificial intelligence), can increase the positive consumer experience in this smart extended reality. In line with RET, the different ways of engagement (Fig. 17.4) in the Metaverse immersion (Yim et al. 2017) are related to the interaction (Nikhashemi et al. 2021), inspiration (Rauschnabel et al. 2019) and satisfaction (Hinsch et al. 2020) of the needs of touch and haptic imagery.

In line with the regulatory engagement theory (Higgins 2006; Higgins and Scholer 2009), the level of engagement in this smart extended reality could derive from level of interaction with other users and smart technologies; the inspiration deriving from the environment able to involve user in acting as in normal life; and finally, from the satisfactions of their needs in a Meta-word such as in the physical word.

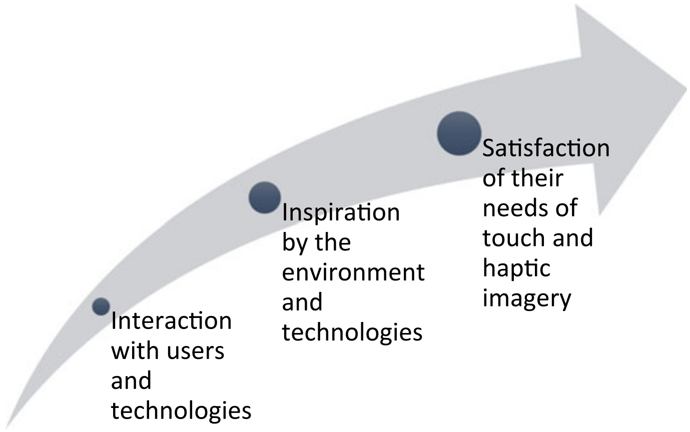


Fig. 17.4 Engagement in the Metaverse

In the Metaverse experience the user experience with the technologies affecting this smart extended reality has a central role on the decision to use the Metaverse to conduct everyday activities as in an everyday life. In this regard, and in line with the regulatory engagement theory, the analysis highlights that the consumer engagement is the principal force able to motivate the Metaverse experience.

17.5 Self-Determination Theory

The self-determination theory (Deci and Ryan 1980; 1985a, b; 1987, 2000, 2012) is based on three principal pillars: autonomy, competence and relatedness. This theory can explain how is essential the need for autonomy for a person in order to self-organize its own life (Deci and Ryan 1985a, 1987). Consequently, the need for competence refers to the ability to be or become effective in each action/interaction with the environment where a person is immersed (Deci and Ryan 1980, 1985a). The need for relatedness (Baumeister and Leary 1995), attend to the necessity to have a feeling connection with other people able to support each action in everyday life. In this regard, the self-determination theory (Deci and Ryan 1980; 1985a, b; 1987, 2000, 2012) could be an excellent support to understanding the principal psychological needs in new smart reality such as Metaverse. People can conduct their everyday activities in a new environment (Metaverse) with new features and characteristics by using technology at their disposal. However, according to this theory, each action is finalized to achieve the need for autonomy, competence and relatedness. People need to self-determine their life (Deci and Ryan 1985a, 1987) to feel control over their choices. For this reason, they need to understand how a smart extended reality could work by discovering its features and potentialities. According to the self-determination theory, people are constantly motivated to grow and acquire new competence skills to satisfy their three psychological needs (Deci and Ryan 2000; 2012). In this context, the Metaverse could be a smart extended reality where people can also satisfy their needs when they cannot conduct the same activities in the physical world.

In fact, in the Metaverse world, people can satisfy their competence needs by learning what they want in digital reality, where they can find several solutions to satisfy this primary need. Then, as the Metaverse is a place enriched by several interactive technologies, people can stay connected with others (relatedness), and preserve their social activity in an unlimited way, thanks to the potentiality of a Metaverse. The self-determination theory (Deci and Ryan 2000; 2012) focuses principally on internal motivation to do something such as the need to gain knowledge or independence (intrinsic motivation), so the Metaverse environment allows people to satisfy their need for autonomy and competence in a smart extended reality. Then, as expressed by self-determination theory (Deci and Ryan 2000; 2012) people need to experience a sense of belonging and attachment to other people, and the Metaverse permits people to achieve a sense of belonging to a community (Baumeister and Leary 1995) by exploiting the potentiality of this world.

To summed up the contribution of the self-determination theory (Deci and Ryan 1980; 1985a, b; 1987, 2000, 2012) to a smart extend reality such as a Metaverse (Fig. 17.5), it is possible to highlight that this environment appears rich of potentiality to satisfy the three basic psychological needs of people. In particular, the Metaverse is a smart environment where people can have at their disposal different way to get new knowledge to satisfy their need of competence. In the Metaverse environment people are also needed to conduct several kind of activities able to support their need of autonomy. And finally, the Metaverse is a place where people can have strong relationships with others, and easily satisfy their need of belonging.

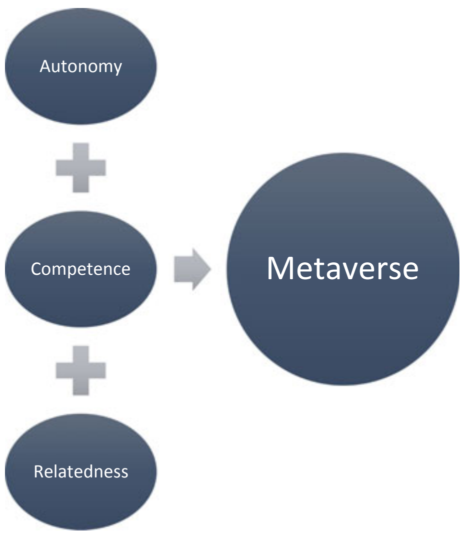


Fig. 17.5 Three psychological needs in the Metaverse

17.6 Discussion and Conclusion

To identify a punctual definition of a smart extended reality such as the Metaverse appear difficult for the paucity of the study in marketing and management able to emphasis the characteristics and benefits and outcomes of this new and not well-defined environment. This chapter aimed to focus the attention on some theories useful and helpful for researcher and practitioners to understand how and from what specific point of view it is possible to understand this new world. More specifically (as represented in Fig. 17.1) the three theories described in this chapter, the affordance theory of technology (Gaver 1991), the regulatory engagement theory (Higgins and Scholer 2009), and the self-determination theory (Deci and Ryan 1980; 1985a; 1985b; 1987; 2000, 2012) could be a valid support to answer how conceptualize a new smart extended reality that is still under investigated. In fact, thanks to the support of the affordance theory of technology (Gaver 1991) this chapter explained how it is possible to get value from specific technology, deriving from the Metaverse world. According to this theory the value from the Metaverse derives from the connection of the elements composing it such as users (real and avatars), the context (technological and digital) where people take place their actions, and the environment (shared, collaborative, multiuser, unlimited) where they conduct their experience.

Then, from the regulatory engagement theory (Higgins and Scholer 2009), this chapter identified how a positive engagement of people during their experience in a specific object or environment could derive from the exploitation of augmented and virtual reality and artificial intelligence, in a Metaverse world. Therefore, in line with the regulatory engagement theory (Higgins and Scholer 2009), this chapter underlined the possible way of involvement in the Metaverse during a shopping journey and exploiting potentialities deriving from the merge of augmented, virtual reality and artificial intelligence, such as interaction, immersion, inspiration and satisfaction. In fact, the level of engagement in a smart extended reality could derive from level of interaction with other users and smart technologies; the inspiration deriving from the environment able to involve user to conduct their actions and regular activities as in normal life; and finally, from the satisfactions of their needs in a Meta-word such as in the physical word.

Finally, the self-determination theory (Deci and Ryan 1980; 1985a; 1985b; 1987; 2000, 2012) identified the intrinsic motivation that led people to experience retailing in the Metaverse-tailing as a fusion of several advanced technologies. According to the self-determination theory (Deci and Ryan 1980; 1985a, b; 1987, 2000, 2012) a smart extend reality (such as the Metaverse) seems to be an environment with high potentiality able to satisfy the three basic psychological needs of people. In particular, the Metaverse appears a smart environment where people can have at their disposal different way to get new knowledge to satisfy their need of competence, and where they can do different activities to satisfy their need of autonomy, and they can also have strong relationship with others, and satisfy easily their need of belonging.

References

Amorim T, Tapparo L, Marranghello N, Silva ARC, Pereira AS (2014) A multiple intelligences theory-based 3D virtual lab environment for digital systems teaching. Proc Comput Sci (29):1413–1422

Arghashi V, Yuksel CA (2022) Interactivity, inspiration, and perceived usefulness! How retailers' AR-apps improve consumer engagement through flow. J Retail Consum Serv 64:102756

Baumeister R, Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull 117:497–529

Binson B (2021) Editorial metaverse and crypto Art during the COVID-19 pandemic. J Urban Cult Res 23:1–2

Caboni F, Hagberg J (2019) Augmented reality in retailing: a review of features, applications and value. Int J Retail Distrib Manage 47(11):1125–1140

Caboni F, Pizzichini L (2022) How the COVID-19 pandemic may accelerate millennials' adoption of augmented reality. Int J Retail Distrib Manage 50(13):95–115

Chen JC (2016) The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in second life. Comput Educ 102:152–171

Choi HS, Kim SH (2017) A content service deployment plan for metaverse museum exhibitions centering on the combination of beacons and HMDs. Int J Inf Manage 37(1):1519–1527

Collins C (2008) Looking to the future: higher education in the metaverse. Educause Rev 43(5):51–63

Cunningham TC (2010) Marching toward the metaverse; strategic communication through the new media. Army Command Gen Staff Coll Fort Leavenworth KS School Adv Mil Stud, VA, USA, Tech. Rep. ADA522953

Dascalu MI, Moldoveanu A, Shudayfat EA (2014) Mixed reality to support new learning paradigms. In: Proceeding 18th International Conference System Theory, Control Computer (ICSTCC), pp 692–697

Davis A, Murphy J, Owens D, Khazanchi D, Zigurs I (2009) Avatars, people, and virtual worlds: foundations for research in metaverses. J Assoc Inf Syst 10(2):90–117

Daz JEM, Salda CAM, Avila CAR (2020) Virtual world as a resource for hybrid education. Int J Emerg Technol Learn 15(15):94–109

Deci EL, Ryan RM (1980) The empirical exploration of intrinsic motivational processes. Adv Exp Soc Psychol 13:39–80

Deci EL, Ryan RM (1985a) Intrinsic motivation and self-determination in human behavior. Plenum Press, New York

Deci EL, Ryan RM (1985b) The general causality orientations scale: self-determination in personality. J Res Pers 19:109–134

Deci EL, Ryan RM (1987) The support of autonomy and the control of behavior. J Pers Soc Psychol 53:1024–1037

Deci EL, Ryan RM (2000) The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychol Inq 11:227–268

Deci EL, Ryan, RM (2012) Self-determination theory. In: Van Lange PAM, Kruglanski W, Higgins ET (eds.), Handbook of theories of social psychology,Sage Publications Ltd., pp 416–436. https://doi.org/10.4135/9781446249215.n21

Dionisio JDN, Burns Iii WG, Gilbert R (2013) 3D virtual worlds and the metaverse: current status and future possibilities. ACM Comput Surv 45(39)

Dolbec YC, Fischer E (2015) Refashioning a field? connected consumers and institutional dynamics in markets. J Consum Res 41(6):1447–1468

Fang Z, Cai L, Wang G (2021) MetaHuman creator the starting point of the metaverse. In: Proceedings of international symposium on computer technology and information

Gadalla E, Keeling K, Abosag I (2013) Metaverse-retail service quality: a future framework for retail service quality in the 3D internet. J Mark Manag 29(13–14):1493–1517

Gaver WW (1991) Technology affordances. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 79–84
Gibson JJ (1977) The theory of affordances. Hilldale, USA 1(2):67–82
Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, New York
Gibson JJ (2014) The theory of affordances (1979). In: The people, place, and space reader. Routledge, pp 90–94
Grewal D, Roggeveen AL, Nordf˙alt J (2017) The future of retailing. J Retail 93:1–6
Grewal D, Hulland J, Kopalle PK, Karahanna E (2020a) The future of technology and marketing: a multidisciplinary perspective. J Acad Mark Sci 48:1–8
Grewal D, Noble SM, Ahlbom C, Nordf˙alt J (2020b) The sales impact of using handheld scanners: evidence from the field. J Mark Res 57(3):527–547
Grewal D, Gauri DK, Das G, Agarwal J, Spence MT (2021) Retailing and emergent technologies. J Bus Res 134:198–202
Guo J, Angelina C, Rolf WT (2011) Virtual wealth protection through virtual money exchange. Electr Comm Res Appl 10(3):313–330
Han J, Yun J, Jang J, Park KR (2010) User-friendly home automation based on 3D virtual world, IEEE Trans Consum Electron 56(3)5606335: 1843–1847
Higgins ET (2006) Value from hedonic experience and engagement. Psychol Rev 113:439–460
Higgins ET, Scholer AA (2009) Engaging the consumer: the science and art of the value creation process. J Consum Psychol 19(2):100–114
Hinsch C, Felix R, Rauschnabel PA (2020) Nostalgia beats the wow-effect: inspiration, awe and meaningful associations in augmented reality marketing. J Retail Consum Serv 53:101987
Huggett J (2020) Virtually real or really virtual: towards a heritage metaverse? Studies Digit Heritage 4(1):1–15
Jaynes C, Seales WB, Calvert K, Fei Z, Grif J (2003) The metaverse: a networked collection of inexpensive, self-con guring, immersive environments. In Proceeding workshop virtual environment, pp 115–124
Jeon JE (2021) The effects of user experience-based design innovativeness on user–metaverse platform channel relationships in South Korea. J Distrib Sci 19(11):81–90
Jeong H, Yi Y, Kim D (2022) An innovative e-commerce platform incorporating metaverse to live commerce. Int J Innov Comput Inf Control 18(1):221–229
Kawaguchi M, Kobayashi T, Yoshitake M (2020) Virtual experiments in metaverse and their applications to collaborative projects: The framework and its significance. Proc Comput Sci (176):2125–2132
Kim J (2021) Advertising in the Metaverse: Research Agenda J Interact Advertising 21(3):141–144
Kim SK, Joo YS, Shin M, Han S, Han, JJ (2013) Virtual world control system using sensed information and adaptation engine. Signal Process, Image Commun 28(2):87–96
Ko E, Jang J (2014) The virtual device managing module of the metaverse assisted living support system. In: Proceedings international conference modeling, simulation visualization methods (MSV) steering committee world congress in computer, engineering and applied Computing, pp 125–126
Kumar S, Chhugani J, Kim C, Kim D, Nguyen A, Dubey P, Bienia C, Kim Y (2008) Second life and the new generation of virtual worlds. Computer 41(9):46–53
Kuvykaitė R, Tarutė A (2015) A critical analysis of consumer engagement dimensionality. Procedia Soc Behav Sci 213:654–658
Kye B, Han N, Kim E, Park Y, Jo S (2021) Educational applications of metaverse: possibilities and limitations. J Educat Eval Health Prof 18(A1)
Laviola E, Gattullo M, Manghisi VM, Fiorentino M, Uva AE (2022) Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing. Int J Adv Manuf Technol 119(3–4):1769–1784
Lee SE, Domina T, MacGillivray M (2011) Exploring consumers' flow experiences in virtual shopping: an exploratory study. Int J Electr Market Retail 4(2–3):165–182

Leenes R (2008) Privacy in the metaverse: regulating a complex social construct in a virtual world. IFIP Int Fed Inf Process 262:95–112

Luse A, Mennecke B, Triplett J (2013) The changing nature of user attitudes toward virtual world technology: a longitudinal study. Comput Hum Behav 29(3):1122–1132

Marmaridis I, Griffith S (2009) Metaverse services: extensible learning with mediated teleporting into 3D environments. In: Lecture notes in business information processing 20 LNBIP, pp 229–239

Messinger PR, Stroulia E, Lyons K, Bone M, Niu RH, Smirnov K, Perelgut S (2009) Virtual worlds past, present, and future: new directions in social computing. Decis Supp Syst 47(3):204–228

Murray JH (2020) Virtual/reality: how to tell the difference. J Vis Cult 19(1)

Nevelsteen KJL (2018) Virtual world, defined from a technological perspective and applied to video games, mixed reality, and the Metaverse. Comput Anim Virtual Worlds 29(1):e1752

Nikhashemi SR, Knight HH, Nusair K, Liat CB (2021) Augmented reality in smart retailing: A (n)(A) symmetric approach to continuous intention to use retail brands' mobile AR apps. J Retail Consum Serv 60:102464

Norman DA (1988) The Psychology of Everyday Things. Basic Books, New York

Norman DA (1999) Affordance, conventions, and design. interactions 6(3):38–43

Owens D, Mitchell A, Khazanchi D, Zigurs I (2011) An empirical investigation of virtual world projects and metaverse technology capabilities. ACM SIGMIS Database: Database Adv Inf Sys 42(1):74–101

Pantano E (2016) Engaging consumer through the storefront: evidences from integrating interactive technologies. J Retail Consum Serv 28:149–154

Pantano E, Gandini A (2017) Exploring the forms of sociality mediated by innovative technologies in retail settings. Comput Hum Behav 77:367–373

Pantano E, Gandini A (2018) Shopping as a "networked experience": an emerging framework in the retail industry. Int J Retail Distrib Manage 46(7):690–704

Papagiannidis S (2008) From 2D to 3D: making the transition from web to metaverse retailing. Cutter IT J 21(9):14–18

Papagiannidis S, See-To E, Bourlakis M (2014) Virtual test-driving: the impact of simulated products on purchase intention. J Retail Consum Serv 21(5):877–887

Park SM, Kim YG (2022) A metaverse: taxonomy, components, applications, and open challenges. IEEE Access 10:4209–4251

Rauschnabel PA, Felix R, Hinsch C (2019) Augmented reality marketing: How mobile AR-apps can improve brands through inspiration. J Retail Consum Serv 49:43–53

Rauschnabel PA, Babin BJ, tom Dieck MC, Krey N, Jung T (2022) What is augmented reality marketing? Its definition, complexity, and future. J Bus Res 142:1140–1150

Rospigliosi PA (2022) Metaverse or Simulacra? Roblox, minecraft, meta and the turn to virtual reality for education, socialisation and work. Interact Learn Environ 30(1):1–3

Rymaszewski M, Au WJ, Wallace M, Winters C, Ondrejka C, Batstone-Cunningham B (2007) Second Life: the official guide. Wiley, Hoboken, NJ, USA

Schaf FM, Müller D, Bruns FW, Pereira CE, Erbe HH (2009) Collaborative learning and engineering workspaces. Annu. Rev. Control 33(2):246 252

Schlemmer E, Trein TD, Cristoffer O (2009) The metaverse: telepresence in 3D avatar-driven digital-virtual worlds. Tic Revista D'innovaci Educativa 2(26):32

Shankar V, Kalyanam K, Setia P, Golmohammadi A, Tirunillai S, Douglass T, Waddoups R (2021) How technology is changing retail. J Retail 97(1):13–27

Sourin A (2017) Case study: shared virtual and augmented environments for creative applications. Springer Briefs Comput Sci 0(9783319540801): 49–64.

Stephenson N (1992) Snow crash; bantam books. New York, NY, USA

Tonéis CN (2011) Puzzles as a creative form of play in metaverse. J Virtual Worlds Res 4(1)

Wiederhold BK (2022) Ready (or Not) player one: initial musings on the metaverse. Cyberpsychol Behav Soc Netw 25(1):1–2

Wright M, Ekeus H, Coyne R, Stewart J, Travlou P, Wright RW, Coyne E, Williams S (2008) Augmented duality: overlapping a metaverse with the real world. In: Proceedings of the international conference on advances in computer entertainment technology, ACE, pp 263–266
Yim MYC, Chu SC, Sauer PL (2017) Is augmented reality technology an effective tool for E-commerce? an interactivity and vividness perspective. J Interact Mark 39:89–103

Chapter 18
Artificial Intelligence and Extended Reality in Luxury Fashion Retail: Analysis and Reflection

Sandra Maria Correia Loureiro

Abstract Retailing service is facing a rapid evolution through the incorporation of Extended Reality (XR) technologies and Artificial Intelligence (AI) algorithms. Within this umbrella, one can consider all real-and-virtual combined environments and interactions generated by computer, including Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR). This chapter brings together research on the experience of luxury fashion consumer retail and AI and XR technologies to identify critical gaps and open avenues for future research. It aims to analyse the incorporation and the potential evolution of AI and XR technologies in luxury fashion retail. The chapter presents the main theoretical concepts and the luxury fashion consumption cycle and offers a discuss about the human and non-human rights, privacy, and ethics at three levels: mega, micro, and nano. Also, the chapter contributes with a discussion about the rights, privacy and ethics involving humans, hybrid humans, and non-humans.

18.1 Introduction

Forbes (2019) estimates that the global market for virtual reality (VR) and augmented reality (AR) in retail will reach USD 1.6 billion by 2025. Innovations such as 3D content can enhance conversion by up to 27.96% on retailer websites (Reydar 2022). Comparing with customers without an immersive retail experience, AR engagement has increased 20% since the beginning of 2020 and the conversion rate increased by 90% (Retailing Customer Experience 2021). Fashion and beauty brands are launching experiences “try before you buy” through filter-based apps. Customers can load up apps in a store to scan for information from AR beacons. One in five U.S.A. customers have used VR retail products in 2020 (XR today 2022). Juniper Research (2022) claims that retail machine learning spend will grow 230% between

S. M. C. Loureiro (✉)
Iscte-Instituto Universitário de Lisboa and Business Research Unit (BRU-IUL) and SOCIUS, Lisbon, Portugal
e-mail: sandramloureiro@netcabo.pt

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*, Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_18

2019 and 2023. Mordor Intelligence (2022) claims that the Artificial Intelligence (AI) in retail market was valued at $1.80 billion in 2019 and is expected to reach $10.90 billion by 2025.

These figures indicate that the retailing service is facing a rapid evolution through the incorporation of immersive extended technologies (XR) and artificial intelligence (AI) algorithms. The term "extended reality" is employed as any technology that extends or creates a new reality by leveraging the 360º space, where technology generated man–machine interactions occur. Withing this umbrella one can consider all real-and-virtual combined environments and interactions generated by computer, including Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR).

XR represents a spectrum of the combination of psychical and virtual reality including Augmented Reality (interactive experience of a real-world environment where objects in the real world are enhanced by computer-generated perceptual information), Mixed Reality and Virtual Reality (the use of high-performance computers and sensory equipment) (Milgram et al. 1994; Matthews et al. 2021).

Real environments regard direct or indirect (e.g., a video display) views of a real scene. Virtual environments are in the other extreme of reality-virtuality continuum (Milgram and Kishino 1994). Virtual environments exhibit a completely computer-generated environment, where unreal objects are displayed on a device, creating Virtual Worlds (e.g., Second Life) and humans are represented by avatars (Flavián et al. 2019). Between the two extremes the Mixed Reality (MR) environments emerge, which are the different points of the continuum where real and virtual objects are merged (Milgram and Kishino 1994). MR creates augmented scene elements in 3D that together with a head-mounted device such as Microsoft Hololens allows the user not only to see virtual objects embedded in the physical world but also to interact with them (Loureiro et al. 2019). MR integrates AR and Augmented Virtuality (AV). AR overlaps virtual objects with actual physical surroundings (Rauschnabel et al. 2017), while AV superimposes real objects on virtual environments (Regenbrecht et al. 2004).

With the advance of technologies, new immersive VR equipment provides a sense of embodiment because users feel that the devices (e.g., HMD, gloves) belong to the human body (Shin 2017). The same occurs with the wearable AR (e.g., AR glasses, AR glasses camera, cognitive systems), where users can extend their natural abilities by enhancing perceptual or motor skills (Yim et al. 2017; Tussyadiah et al. 2017; Javornik et al. 2021). The embodiment can go so far as the full integration of devices in the human body (e.g., microchips, smart contact lens, nanotechnology), which can complement or improve the cognitive abilities of users (Wilson 2002; Tussyadiah 2014). For instance, with smart contact lens it is possible to project visual information directly into the retina of the eye. Therefore, due to the new technological developments, Flavián et al. (2019) propose an adjustment of the reality-virtuality continuum by introducing the pure mixed reality, where virtual elements are rendered in such a way that they cannot be distinguishable from the physical environment (Collins et al. 2017). Users can interact simultaneously and in real-time with real and virtual objects.

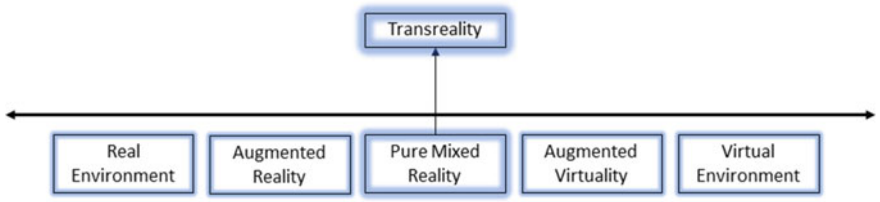


Fig. 18.1 Extended reality-virtuality continuum

As immersion technologies advance towards a more embodied incorporation of devices into the human body that intensify the senses and emotions, change the physical body, or enhance cognitive abilities, the sense of being elsewhere (Collins et al. 2017) will be more effective than it is today. Likewise, as humans incorporate more immersive technologies, their physical bodies and minds will be more capable to modify the position, orientation, or features of virtual and real objects. In those future days we can claim that enhanced humans will experience transreal environments. These environments will go beyond what today we recognize as pure mixed reality (see Fig. 18.1). Transreality represents here the use of nanotechnology (incorporation of nano-devices in a scale of 100 billionths of a meter) to extrapolate and reshape the human body capabilities in terms of sense, emotions, memory, and physical features allowing the experience of unique realities.

AR and VR are combined with robots in the fashion industry to enhance the interaction between humans and robots (Sanderson 2018). Industrial robot programming systems can be assisted by AR. This programing process uses augmented graphics (e.g., the visualizations of the user-defined points and paths) directly onto the surfaces of workpieces (e.g., Ong et al. 2020; Makhataeva and Varol 2020). Robots are also coming to the frontline, acting as frontline assistants, and interacting with customers. Robots can have artificial intelligence (AI) algorithms incorporated and learn how to deal with customers (Belanche et al. 2020). AI means the "programs, algorithms, systems and machines that demonstrate intelligence" (Shankar 2018, p. vi) and represents the "use of computational machinery to emulate capabilities inherent in humans, such as doing physical or mechanical tasks, thinking, and feeling" (Huang and Rust 2021, p. 32).

Academics create different typologies for the levels of AI (Mende et al. 2019; Loureiro et al. 2021a, b). Davenport and Kirby (2016) suggest task automation (standardized or rule-based applications) and context awareness (requires algorithms to "learn how to learn" and extend beyond their initial programming). Huang and Rust (2018) consider four AI: mechanical, analytical, intuitive, and emphatic. Mechanical is the ability to automatically repeat tasks and routines without any help from humans. They are extremely consistent in performing tasks without the human fatigue aspect. As examples of jobs that can be performed by Mechanical AI is the operation of retail salespersons or waiters/waitresses. Analytical AI can perform complex, yet systematic and predictable tasks. Machine learning and data analytics are the major analytical AI applications. Thus, these are machines able to process and synthesize

Table 18.1 Typology of levels of intelligence

Typology				Authors
Task automation		Context awareness		Davenport and Kirby (2016)
Mechanical AI	Analytical AI	Intuitive AI	Empathetic AI	Huang and Rust (2018)
Narrow AI		General AI		Kaplan and Haenlein (2019)

large amounts of data (big data) and learn from that information. For instance, the large number of reviews from social media can be analysed by Analytical AI. Intuitive AI can think creatively and is able to adjust to novel situations (Sternberg 2005). As Huang and Rust (2018, p. 159) claim that "tasks that are complex, creative, chaotic, holistic, experiential, and contextual require intuitive intelligence". For example, intuitive AI can operate as creators in advertising or creative problem-solving. Empathetic AI can understand emotions and interact emotionally (Goleman 1996). Sophia, the human-like AI, from Hanson Robotics and Replika are first examples of what such Empathetic AI could be.

Kaplan and Haenlein (2019, p. 15) consider AI as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation". They propose Narrow AI versus General AI, where the first is focused on a specific domain and the second can extend into new domains. Thus, Narrow AI is somewhat associated with mechanical and analytical intelligences, whereas General AI with intuitive and empathetic intelligences (Reese 2018) (see Table 18.1). Davenport et al. (2019) argue that the level of intelligence can be conceptualized as a continuum.

Although luxury fashion retail managers are becoming more open for incorporating AI and immersive technologies, to date academics have not been prolific in developing research that explores the use of these technologies in luxury fashion retail (Chung et al. 2020). Only a few studies attempt to analyze AI and immersive technologies in fashion retail, but they are not specialized in luxury consumption (e.g., Watson et al. 2020; Tupikovskaja-Omovie and Tyler 2020; Baytar et al. 2020; Hamouda, 2021; Loureiro et al. 2021a, b). Academics tend to focus primarily on studying specific technologies and concepts, without a broader perspective of the experience that customers are or will be living in their fashion consumption. For instance, understanding how luxury fashion retail can provide a personalized e-service through AI agents (chatbots). Other studies are concerned in developing digital shopping platforms for fashion shopping where customers meet their expectations, using eye tracking fashion m-retail (Tupikovskaja-Omovie and Tyler 2020), augmented reality for virtual try-on (Baytar et al. 2020), or explore the customer experience during the interaction between customers and the mobile application (Watson et al. 2020; Hamouda 2021). Likewise, exploring how the experience in a virtual shoe store with different music tone stimulates the consumers' cognitive and affective state (Loureiro et al. 2021a, b).

Understanding how the increasingly complex technologies affect customer experience is also a priority suggested by the Marketing Science Institute (MSI) (MSI

2021). Therefore, this chapter intends to bring together research on the experience of luxury fashion consumer retail and AI and immersive technologies to identify critical gaps and open avenues for future research. This chapter aims to analyze the incorporation and the potential evolution of AI and immersive technologies in luxury fashion retail. The main research questions are:

- How do AI and immersive technologies affect luxury fashion consumption cycle currently?
- How will the potential evolution of the AI and immersive technologies influence luxury fashion retail?
- What might be the concerns in terms of human and non-human rights, privacy an ethics associated with the context of luxury fashion?
- What might be questions for future research?

To do so, the chapter first presents the main theoretical concepts and the luxury fashion consumption cycle. Then the chapter offers a discuss about the human and non-human rights, privacy, and ethics at three levels, mega, micro, and nano. Finally, the chapter provides the research agenda. This chapter will contribute as a first reflection on AI and immersive technologies of actual and future evolution on luxury fashion retailing, proposes the luxury fashion consumption cycle model for AI and immersive technologies, and suggests research questions for the future. The chapter also contributes with a discussion about the rights, privacy and ethics involving humans, hybrid humans, and non-humans.

18.2 Retail Customer Experience

Customer experience is a multidimensional concept that represents emotional, behavioral, cognitive, social, and sensory responses of consumers to the retail stores offered during the purchase journey (Schmitt et al. 2015; Verhoef et al. 2009). Diverse stimuli during the purchase journey shape consumer experience. Those stimuli come from the atmospheric cues of the physical or online store (e.g., layout, design, colour scheme, scent, music background), frontline employees, past experience, and marketing communication efforts (e.g., marketing campaign, social media).

Layout and design mean how the store is set up for the customer flow. When the layout is too complex, it helps to retain customers in the store (physical or online), but it can also annoy them and invite them to leave without buying. Warm colours are more associated with excitement, energy, and movement than cool colours (Crowley 1993; Labrecque et al. 2013). Scent represents aromas or smells that come naturally from the store (as in bread store) or purposefully (e.g., use a fragrance to represent a brand) to induce pleasure in consumers (Spangenberg et al. 2006). Music is also an ambient element that influences emotions and behaviors (e.g., Grewal et al. 2003; van Rompay et al. 2012). Frontline employees provide the service, contribute to a more personalized service, and to develop the interaction between the retail brand and the customer. They are important elements in the customer experience because they can induce emotions and behaviours, such as pleasure, enjoyment, commitment, and the return to the store (Palmatier et al. 2006).

Past experience influences the appraisal of the current experience because it creates expectations (Verhoef and Van Doorn 2008). When consumers have a good memory of the previous experience, this can help them to recommend the store (physical or online) to others or return more often. However, a negative experience is difficult to overcome and can lead the consumer to complain or stop going to the store (Verhoef and Van Doorn 2008). As past experience, marketing communication also forms the expectations about the experience. Marketing communication can take many forms from in-store communications to advertising campaigns or social media. In-store communications represent all type of signage that guide customers in the store (Van Nierop et al. 2011). Advertising campaigns are created by the retailers to promote their products. Although retailers can also use social media to promote their products, this medium is mainly used by customers to comment on the products, by recommending or complaining (Van Doorn et al. 2010). Social media can be a way to engage the customer (Hollebeek et al. 2014).

Although distinct from experience, the concept of engagement is closely associated with it. A good product or brand experience can contribute to engage customers. By customer engagement Van Doorn et al. (2010, p. 253) claim that it is "the customer's behavioral manifestation toward a brand or firm, beyond purchase, resulting from motivational drivers". From the relational perspective, Kumar and Pansari (2016, p. 498) define engagement "as the attitude, behavior, the level of connectedness (1) among customers, (2) between customers and employees, and (3) of customers and employees within a firm". Customer engagement has a multi-dimensional nature, where the core meaning refers to interaction (e.g., repeat purchase, exchange ideas), emotional connection, and cognitive focus towards the object (i.e., retail, product/brand).

18.3 Luxury Fashion Consumption Cycle

Luxury fashion brands have global reputations, associated with innovation and tradition and encouragement of the customer's desire (Tynan et al. 2010). Hermes, Chanel, and Louis Vuitton are commonly ranked in luxury brand indexes (e.g., Prestige Brand Ranking 2021; Interbrand 2020). Ko et al. (2019) claim that a luxury brand is perceived by customers as being high quality, authentic, having a prestigious image, commanding premium pricing, and inspiring a deep connection. Fashion brands as Hermes, Chanel, and Louis Vuitton, Gucci, Prada, Burberry, or Dior have such characteristics. In the luxury brand industry two terms are often employed: *Haute Couture* (high fashion) and *Prêt-à-porter* (ready-to-wear). The first refers to one-of-a-kind piece, completely personalized, unique, with extreme attention to detail. As for the *Prêt-à-porter*, these are pieces of high-end garments that are available at physical or online sophisticated stores. Although high fashion is extremely expensive and is affordable only to a very few, ready-to-wear has premium prices, but is more accessible to customers.

The high fashion (*Haute Couture*) pieces are all handcrafted, bespoke and require a lot of skill and time. *Haute Couture* customers do not buy from retail stores, they belong to a very exclusive high wealth percentile, so the pieces are made exclusively for them, and they have direct access to the couture designer's stylist of the Couture House. Ready-to-wear stores are the core focus of this discussion. Thus, are these brands using XR and AI in their retail? The answer is yes. Gradually luxury brands are introducing these technologies, particularly in the ready-to-wear context.

For example, Dior Virtual Try-on is an app that allows customers to try on sunglasses and headbands via AR on Instagram, from the comfort of their own smartphones. They also launched a virtual sneaker try-on with Snapchat. Dior Insider is an AI chatbot that promotes the brand's beauty to help consumers keep up to date. Dior Insider begins by greeting the user by name and asking if he/she is a Dior privileged customer. If the user answers no, Dior Inside says "Not yet… it's just a matter of time" with a wink emoji. The VR headset—Dior Eyes—is equipped with high-definition image resolution and integrated holophonic audio, creating a 3D immersion into the backstage world of a fashion show.

In a partnership with digital clothing marketplace DressX, designer Clara Daguin wears a digital version of her Jacquard in the Paris Haute Couture Week—Fall 2021. The designer Damara Ingles presents Sutu wearing looks originally created for the VR Paris fashion show 2020, claiming that the future of fashion is a symbiose between biological bodies and virtual environments (Voguebusiness 2021).

Other example comes from Burberry 'ARKit', which is the Apple's augmented reality for iOS (mobile operating system) that allows customers to digitally redecorate their surroundings with Burberry-inspired drawings by artist Danny Sangra. Coach installed VR headsets in stores in ten malls across the United States of America provide customers with full access to its latest runway show. Louis Vuitton initially took more time than other brands to adopt these technologies, but now it demonstrates AR technology-infused sneakers and accessories and presents a new textile that allows for flexible displays of high-resolution imagery in bags, which they called "Canvas of the Future".

What will the future of luxury retailing bring? The pandemic situation of COVID-19 has given us a glimpse by opening the consumers' mind to accept technologies that allow them to be closer to the luxury fashion brands, even without a traditional physical touch point in a physical retail store. Luxury fashion industry is known for being creative and innovative, making the fashion experience dynamic and iterative.

Dynamic, because *Haute Couture* houses create a collection every winter and summer season (a very creative industry whose innovation takes place twice a year). *Haute Couture* tends to inspire or even define the fashion trends of the ready-to-wear of the same luxury fashion brands. Iterative due to the cycle that flows from before purchase, effective purchase and after purchase. The cycle process (see Fig. 18.2) shapes the luxury fashion consumer experience over time.

In the future, the three major phases of the consumption cycle are expected to evolve to customer-brand relationships with multiple interactions between humans,

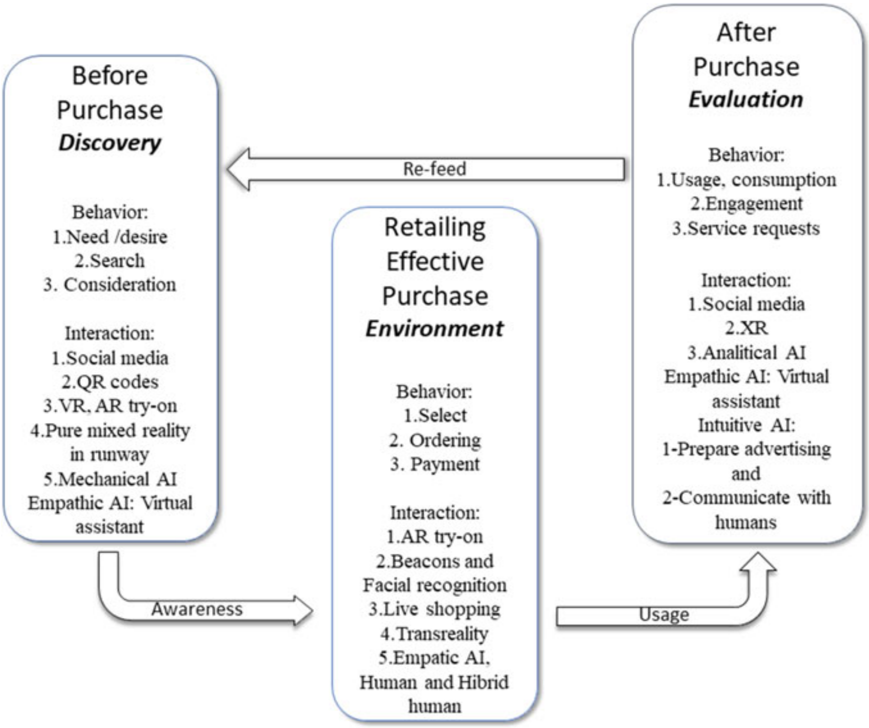


Fig. 18.2 Luxury fashion consumption cycle model for AI and immersive technologies

machines, and hybrids. Hybrid means beings that are humans but have incorporated technologies (as nanotechnologies) to enhance their capabilities, what can also be called as transhuman (e.g., Bostrom 2005; Kumar et al. 2016; Henkel et al. 2020). Thus, luxury fashion customers in subsequent years may constitute a heterogeneous market of human, hybrid, and non-human individuals. The fashion luxury providers can also be composed by the same multiplicity of individuals.

In this vein, the creative stylist of a Couture House could become a hybrid or a non-human (for instance, an AI agent that has evolved into a more creative and intuitive way). In the ready-to-wear manufacturing process, both human and non-human employees can cooperate. Some hybrid humans with enhanced physical and cognitive abilities can, alone or in interdependence with AI algorithms, design more personalized clothing or be more innovative in creating new green fabrics and materials. In *Haute Couture* the tailor-made outfit and experience can be performed by highly evolved AI agents. These agents (robots or not) can operate with large amount of data, knowledge psychological and physical aspects of human beings and provide unique experiences alone or with other human employees.

18.3.1 Before Purchase

This phase (see Fig. 18.2) aggregates the interaction activities between consumers and luxury fashion brands before the act of purchase. Here customers recognize that they desire a luxury piece (Boujbel and d'Astous 2015). They search and consider alternatives. Luxury consumption is associated with hedonic pleasure (Hagtvedt and Patrick 2009), the wish to create a distinction from the mainstream and to belong to a niche group of consumers (Tungate 2014). Luxury fashion brands instill in consumers the feeling of uniqueness, excitement, upper-scale, and distinctiveness (Prentice and Loureiro 2018).

The use of VR to communicate brand identity may enhance the attractiveness of the brand. VR creates a psychological state of isolation from the real world leading to a flow in which the fashion customer is completely taken by the fashion brand. As this technology technically evolves, luxury consumers can have the whole experience of using a fashion piece, without actually having it. They may not only visualize the colour and shape, but also touch, smell, and use all senses, depending on what the brand wishes to provide (Petit et al. 2019). A luxury fashion consumer can have a pre-purchase experience of different luxury brands prepared specifically for him/her. For instance, the runway show can be prepared in a sequence of pieces that can better fit the style and personality of the fashion consumer. In 2021, a VR movie showing a runway is already possible but with the current state of the technology a customization is not possible, neither is the use of the five senses without a CAVE (Cave Automatic Virtual Environment—a cubic room where stereoscopic images are projected) (Loureiro et al. 2019).

A few luxury brands—as Ralph Lauren and Louis Vuitton—are using QR codes for their advertising campaign. QR codes can connect fashion consumers with AR experience for a more insightful, interactive, and playful demonstration of the fashion pieces and trends (Wedel et al. 2020).

The use of virtual try-on technology (AR apps)—as for instance in Dior makeup or Gucci shoes—provide a luxury experience before any purchase decision that can engage the consumer, encouraging and leading them to the effective purchase online or offline (Spence et al. 2014). AR apps for mobile devices can give information about the characteristics of the products, such as colour, texture, size, price (Gäthke 2020). However, the smell of materials or the touch of products to understand their texture is still not possible. Even so, the human brain can cognitively process stimuli that are not experienced through memories of past experiences (e.g., memories of smell and touch of similar fashion pieces), that is, mental imagery (Niedenthal et al. 2005; Petit et al. 2019).

The theory of embodied cognition claims that all cognitive processes are due to both bodily states and specific processing systems of the brain's sensory modality (Barsalou 2008). Thus, when luxury fashion consumers experience the fashion piece in a real setting, they mainly activate the cognitive bodily states process. The brain integrates the perceptual, introspective, and motor state resulting from the stimuli of smelling and touching, seeing, and wearing the garment in the real world. The multisensory representations are stored in the fashion consumers' mind (Petit et al. 2016). Then, when fashion consumers use VR or AR try-on environments without being able to use all the senses, the specific processing systems are activated with the memories created by past experience of similar fashion pieces. For example, the consumers' visualization of shoes can incorporate mental simulations of past experiences using shoes in them and producing spontaneous sensations similar to actually touching and wearing the shoes (Okajima et al. 2016). In the future, with the evolution of technology, the immersive experience will be increasingly bespoke, and the five senses will be experienced in any place and at any time.

The pure mixed reality is a way to provide an immersive experience of fashion pieces (Flavián et al. 2019), which makes the virtual elements so embedded in the physical environment that fashion consumers cannot see the difference between real and virtual. This can be an important advance for luxury brands to promote their luxury pieces. The virtual runway can be more creative, providing an experience combining real and virtual pieces tailored to small groups of fashion customers.

Analytical AI integrated in a system with Intuitive AI can cooperate to use the big data of customers and prepare creative virtual and personalized VR environment. Social media, eventually with new forms, is an example of where empathic AI could be incorporated (as if it were a human being) and where AI can act as the steward (e.g., a sophisticated chatbot or an AI avatar with a deep knowledge of the identity of the mother house) of the luxury brand. One may imagine a scenario where a certain fashion luxury brand has an AI empathic steward for each customer with a face, body, voice, gender that can represent the identity of the brand. The luxury fashion customer feels identified and emotionally involved with the specific steward developing an intimate relationship of mutual trust and interdependence (Fournier 1998; Sampson and Chase 2020). The same steward is also present in VR or pure mixed reality personalized immersive environment, accompanying, and giving suggestions to the customer in a runway prepared specifically for him/her. In this context, the luxury clothes can be co-created with the customer and experienced even before the physical existence. This phase creates awareness for the customer to make the decision to the effective purchase.

18.3.2 Effective Purchase

This phase encompasses the interactions when the luxury fashion customer is effectively purchasing. Traditionally this phase is connected to the selection, ordering and payment inside a real store or using online platforms (De Keyser et al. 2019; Bromuri et al. 2021). Various research on atmospheric cues have been conducted to understand how the environment stimuli—such as design, layout, music, scent, and colour—can affect purchase (Roschk et al. 2017; Bolton et al. 2018). In 2021, physical stores are using mirrors with AR to help customers in the selection process by visualizing themselves with clothes without actually putting them on. Online stores also present environment stimuli through the luxury brand website, which uses atmospheric cues as design, layout, colours, links (Roschk et al. 2017), chatbots and try-on. Luxury brands also use chatbots in Facebook's Messenger or WhatsApp. This chatbots (AI virtual assistants)—for instance Burberry has Lola—guide the customer in the online store.

The actual Beacons—small wireless devices that transmit a continuous signal; the cloud server delivers specific content to the fashion smartphone—and facial recognition—scanning thousands of landmarks on a person's face to identify their unique 'face imprint'—allow to track customers in the store and provide specific information about the products that may interest them, as does the app of Macy's luxury malls.

However, the live shopping (combining video livestreaming, e-commerce, and social media; simultaneously recorded and broadcast in real-time) that Chinese luxury fashion customers experience with Bilibili is taking the online experience to a more vivid reality. Louis Vuitton and Marc Jacobs are already taking advantage of this platform. Live shopping allows for an interactive in-store experience that creates positive emotions and increases engagement (Hollebeek et al. 2014; Kumar and Pansari 2016) with the luxury brand. The Pandemic situation of Covid-19 accelerated the use of live shopping for luxury brands.

Luxury fashion brands are recognized for innovation, exclusivity and bespoke pieces and experiences. Thus, to be competitive and offer innovative and tailored pieces and experiences, luxury fashion brands will consider the evolution of technology to make the fashion consumer experience transreal environments. Transreal environments will be so immersive that the customer will not be able to distinguish whether they are in an online or offline environment, but it will be a unique, innovative, and personalized experience. Luxury fashion customers who desire to incorporate nanotechnology devices in their bodies can be in direct interaction with the fashion brand without the interface of a laptop, table or mobile. The image of the new pieces and the luxury environment will emerge in their mind, and they (brand and customer) will interact in real-time.

The fashion customer can have a personal fashion stylist, which will be a human or an AI algorithm. Customers can think and feel that they are actually wearing the piece because the nano device can operate in the mind creating the feeling of real usage. In this high-tech environment, the customer can also have the option of purchasing the physical part or the image of using the virtual part. The sensation will be very similar in the customer's mind, but in the case of the physical piece customers can own the object (the luxury piece) and in the case of the virtual piece, they can have the sensation of using it for a period of time. In both situations the feeling of wearing the piece will be very real in the consumer's mind.

In this new market, customers are not only humans. They can also be hybrid humans or AI agents. AI agents that evolve to become intuitive and empathic can eventually have their jobs and be introduced in the society as any other human. Thus, AI emphatic agents may develop a desire for fashion pieces. They can even be influencers in social media (as humans are digital influencers today) or even be virtual influencers. By virtual influencer means AI (without virtual body as an avatar) with the ability to change opinions and behaviours in a transreal environment.

In the future, empathic AI robots can assist customers in physical and more traditional stores, particularly in the case of ready-to-wear (prêt-à-porter) luxury stores (Belanche et al. 2020) in a more tailored way. They may empathize with regular customers and understand their desires and style, giving them suggestions, advice. More select and exclusive customers can have the same steward, designed for each individual customer, as in the previous phase. Thus, very exclusive fashion customers can benefit from having a steward in all phases of the purchasing process and use a unique virtual room to select and effectively purchase the fashion products. Nevertheless, real private rooms can still exist for a strict number of customers that wish to try on clothes and can be assisted by humans and/or robots. AI can operate as systems with or without a vocalization similar to a human voice (e.g., Amazon Echo, Apple Siri) or as robots. Robots with AI can work side by side with humans. AI robots are evolving to learn how to walk and master complex human skills (Fevre et al. 2019; Singh and Bera 2020), in sum, to have human-like movements and behaviours. With the effective purchase phase, the luxury consumer will be able to have an experience of using the purchased luxury piece.

18.3.3 After Purchase

This phase is traditionally associated with the use of the fashion products, service support regarding, for instance, guarantee and complaints, and the engagement with the brand. Customer engagement is a complex construct representing the consumers' interaction with the focal object (e.g., the luxury fashion brand) (Van Doorn et al. 2010; Bilro and Loureiro 2020). The engagement process can create mutual interest and behaviours between the luxury brand and the customer through word-of-mouth, giving recommendations, writing reviews in social media, or buying more often products of the same brand (Kumar and Pansari 2016).

At this phase, luxury fashion customers use, touch, and feel the luxury fashion products in their daily life or use exclusive pieces at special moments, as parties, galas. For instance, celebrities parade models at the Oscars' runway and other public figures use models bought, given, or borrowed by fashion luxury brands at various events. These people with social influence communicate the brand by attending public events but they can also be digital influencers by making comments via traditional social networks. An intuitive and empathic AI can also be an influencer that is able to adapt to different circumstances and eventually have a virtual body to exhibit luxury fashion products. In a transreal environment and using their virtual body (e.g., AI avatar) they become a virtual influencer.

Therefore, AI is a relevant tool in producing a tailored engagement marketing approach. AI can enable real-time learning, leading to an increase of value for customers, customer retention and growing competitive advantage (Kumar et al. 2019). AI systems are able to develop persuasive communication with human employees, capture the essence of the communication to assist in promoting the service and formulate questions that contribute to solving problems (e.g., Aicardi et al. 2018).

This phase opens the window to the re-feed. When luxury fashion customers have the experience of using the fashion piece, they will evaluate such usage experience. If the evaluation is positive, customers recommend the brand to others and are more willing to search for more information about the brand with the intention of purchase it again in the future. If the evaluation is negative, fashion customers will tend to provide negative information about the brand and will search for other brands.

18.4 Human and Non-human Rights, Privacy and Ethics

The AI and immersive technologies bring to the public discussion issues related to human and non-human rights, privacy, and ethics. This discussion can be conducted at three levels: mega, micro, and nano (see Fig. 18.3).

18.4.1 Mega Level

This level comprises governments, legislators, and international Institutions. International organizations, governments, and legislators weave the laws and rights that manage the life of societies at a global level or referring to each country. Intergovernmental organizations as United Nations that created human rights will also have to pronounce on the rights of non-humans.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) recognizes the need for international and national policies and regulatory frameworks to ensure that the emerging technologies do not harm Humanity as a whole (UNESCO 2021). They claim that AI should exist for the greater interest of humankind and not the opposite.

Legislators and ethicists worldwide have begun to develop legal norms and standards to regulate the emerging technologies, particularly AI algorithms. Some examples are the Montreal Declaration for Responsible AI, the Asilomar AI Principles, the AI 4 People's principles for AI ethics, the two High-Level Expert Groups on AI's reports (European Commission), on ethics as well as governance of AI, the House of Lords Artificial Intelligence Committee, and the German Ethics Code for Automated and Connected Driving, which entail important aspects of ethical issues related to AI.

The Montreal Declaration emphasizes ten principles (Montreal Declaration 2018), namely, artificial intelligence systems must (1) permit the growth of the well-being of all sentient beings, (2) respect the people's autonomy, (3) protect the people's privacy and intimacy, (4) be compatible with maintaining the bonds of solidarity among people and generations, (5) meet intelligibility, justifiability, and accessibility criteria, and must be subjected to democratic scrutiny, debate, and control, (6) contribute to

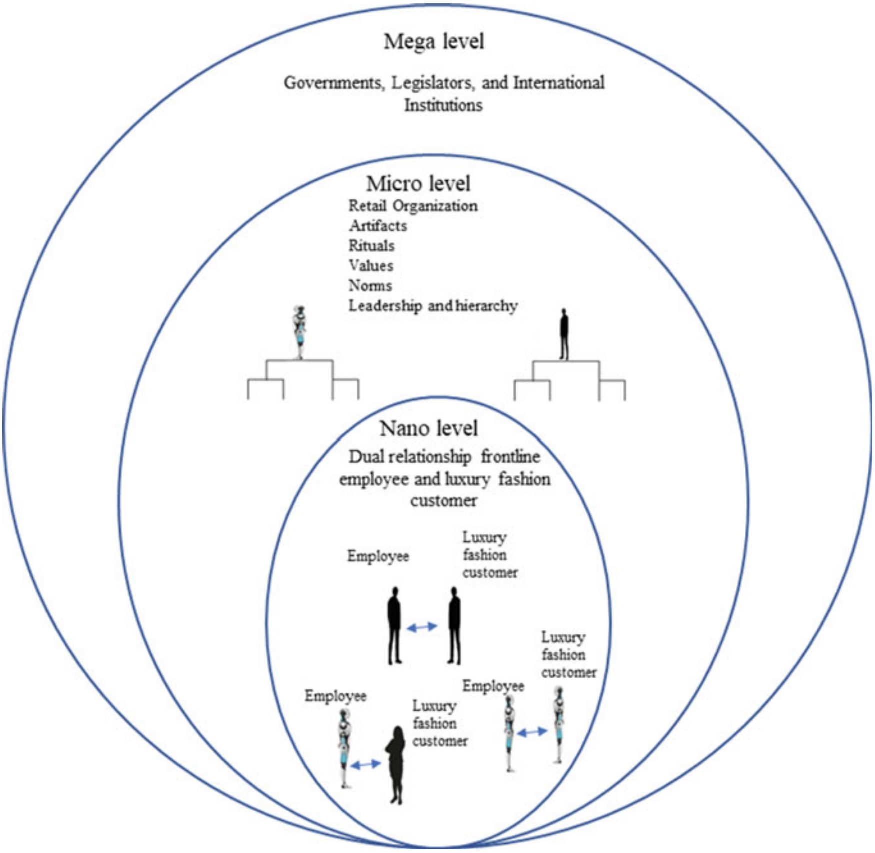


Fig. 18.3 Human and non-human levels of rights, privacy, and ethics

the creation of a just and equitable society, (7) be compatible with maintaining social and cultural diversity and not to restrict the scope of lifestyle choices or personal experiences, (8) every person must exercise caution by anticipating the adverse consequences of AI, (9) not to contribute to lessen the responsibility of human beings when decisions must be made, (10) to be carried out so as to ensure a strong environmental sustainability of the planet.

Aligned the Asilomar AI's Ethics and Values and the AI 4 People's principles for AI ethics mention the need for secure, equitable, fair, transparent, private, and responsible AI systems, that preserve human privacy (Floridi et al. 2018; Futureoflife 2019). AI systems should be designed so that their goals and behaviours can be aligned with human values. Humans are the ones that decide what to delegate AI systems and these systems should respect and improve, rather than subvert, the social and civic processes on which the health of society depends. The European Union states that AI systems must be controlled by humans, AI systems must empower human beings to make more informed decisions, have transparency, respect for privacy and data protection and adequate data governance mechanisms (Europa 2019). The U.K. parliament and the German Ethics Code for Automated and Connected Driving are other voices expressing the importance for transparency, privacy, and human empowerment.

All these committees point out human empowerment and treat AI systems as ways to support human, what is quite reasonable for mechanical and analytical AI systems, that is the narrow AI, as we have today. Yet, AI systems and agents are designed to learn and evolve. They will become intuitive and empathic. AI agents will be bold in terms of creativity, similar or even transcend human regarding cognitive and even emotional states. In this situation can we humans consider them as slaves or assistants of human beings, or they will become a non-human species? In this situation what will be their rights? Similar to human-beings?

18.4.2 Micro Level

Organizational culture encompasses artifacts (e.g., stories, rituals, arrangements, a flag, colors having a strong symbolic means), shared norms and shared values (Deshpande and Farley 2004). Culture influences how the organization adapts to changes and contributes to motivate employees. The incorporation of technologies is a factor whose incorporation requires changes in the organization. The acceptance of technologies as AI in luxury fashion retailing organizations involves the recognition that the benefits are meaningful and the risks can be minimized, prevented, or controlled through risk management (e.g., insurance and redressing). Fashion retail organization ethics comprises organizational values and norms that guide individuals' behavior (Ferrell et al. 2019). Thus, AI systems should not only incorporate the norms and values stablished by supra national organizations, but also the specific ones that each retail organization has. Although rituals, stories and other artifacts can be incorporated in the AI algorithms, this is not a necessary condition since AI can easily learn them all over the time.

In luxury fashion retail, managers need to discuss the decision-making process and the hierarchical structure in an environment of high technology with AI. The AI decision-making process can include feelings and emotion-based personalization similar to human beings (Huang 2017), but the ai ethical choices will emerge in behaviours rather than in intentions (Etzioni and Etzioni 2016). so, ethics should be integrated in the design of the ai algorithm (Martin et al. 2019).

With ethic principles implemented, AI can be allowed to take diverse decisions through predictive analytics, implement operations and associated with retail frontline service (Kaplan and Haenlein 2019). But can luxury fashion retailer leaders let AI make strategic decisions? If so, that means humans will transfer the decision power to the AI. If not, when in the future AI systems transcend humans cognitively and emotionally, these systems will not take the leadership, whether humans want it or not? The use of AI nanotechnology to enhance human cognitive and emotional capabilities can create hybrid humans whose potentials transcend those of current humans. Are these hybrid beings to lead in the future? Therefore, in the hierarchical structure humans can be led by AI systems or by hybrid beings. Can we avoid it today? Or it will be an inevitability due to the main assumption of AI, that is, its ability to continuously learn and adapt.

The luxury fashion industry is well-known for its fast innovation and creativity. This industry creates trends and is open-minded. But can AI systems transcend the creativity of humans? At least AI systems can cooperate in duality with human creativity giving the possibility to identify faster than humans, certain particularities—the new *zeitgeist*—in the street and social media that boost the innovation of luxury fashion.

While AI systems can handle a large amount of data, they need it to be relevant and accurate to be effective in the predictions they make and tracking new trends. This point can collide with the privacy of the fashion consumers. Thus, the level of privacy imposed by the luxury fashion retail brand or the level of trust that the consumer places in the brand may restrict the personal data that is provided. Personal data means sociodemographic variables (e.g., age, gender, income, facial expressions), but also more sensitive data such as photos, medication, or political preference. Sensitive personal data can come from tracking retail consumers when they look for certain clothes in physical stores or online. For instance, the use of augmented reality apps can save information about luxury fashion preferences. Virtual reality technology can also incorporate eye tracking to understand which specific points of the store or luxury clothing customers are looking at or take longer to observe. The virtual runaway—often used during the Pandemic situation—and the traditional runaway—using, for instance, eye tracking tool associated with augmented reality—can provide useful data for AI systems. The use of AI algorithms to analyse street wear can add to the other more traditional information about fashion consumers (particularly some *avant gard* and well-known streets in different parts of the world). Yet, retailer and luxury fashion brands need to create guidelines about how all this information can be used.

AI systems are powered through different datasets (De Bruyn et al. 2020) and depending on the trustworthiness of the information accessed, it will be possible to have a better or lesser forecast of the trend, thus, be a more accurate means to reduce bias, and this lies on the quality and quantitative of the data. The policy to restrict information due to the privacy of luxury fashion consumers can decrease the accuracy of the AI analytical predictions.

Luxury fashion retailers should understand their customers sensitivity towards the donation and use of their personal data. The knowledge about the customers and the general guidelines of the international organizations help to trace the policy of ethics and privacy. With the evolution of the AI systems and other technologies the policy of ethics needs to be revised from time to time.

18.4.3 Nano Level

Nano level represents the dual relationship between frontline employee and luxury fashion consumer. A human frontline employee must be trained to perform his/her tasks in accordance with the retail's organizational culture and identity. He/she has labour rights and duties in accordance with the laws in force in the countries where the retail is implemented and in accordance with the general policy of the fashion luxury brand. However, frontline employees are not just human, they are also AI agents (e.g., chatbots, robots in stores). These AI agents are evolving in knowledge and emotionally, and eventually they will become more aware of its existence and one day these non-human high evolved agents will begin to claim their rights. Eventually they could become consumers. At the same time hybrid humans—those who incorporate nano-technologic devices or adopt prosthetics that enhance their capabilities—can also perform the double role of being employees and customers.

In this situation we will witness a proliferation of dual interactions between human-AI agent or human–human with different levels of technology incorporated. Luxury fashion retail will operate in multiple and interrelated places that will go beyond the well-known physical or online stores. The transreal environments integrating multiple humans, non-humans and AI agents will reshape each individual values and ethics towards life and purchase behaviour. Individuals will tend to analyse in real time what are the benefits and risks of purchasing a particular luxury piece, which will increase the front-line employee's responsibility to persuade and trust customers.

In transreal environments, it will be difficult to distinguish what is real (in the sense of what we know today) and what is not. In these circumstances, it will be the luxury brand's ethics to prevail in order to generate confidence that their data will not be misused, since in this environment, privacy will be difficult to achieve.

18.5 Conclusion

XR and AI technologies are bringing a revolution to the fashion retailing business. VR and AR technologies will evolve to become easier to be used, with updated oculus or other devices to immerse in VR experiences or interfaces that are not smartphone or tablets for AR.

XR technologies can evolve to a pure mixed reality, in a first stage, and then to transreality. In the pure mixed reality luxury fashion consumers cannot distinguish from virtual and physical elements. However, in transreality the immersive environment is leveraged to the extreme, where beings incorporating nanotechnologies in their bodies can interact simultaneously in real and virtual environments without any external equipment and without differentiating between environments. The dynamic interexchange in different environments is not only possible to hybrid humans. AI algorithms can also have access to such environments. They can present themselves as AI avatars in such way that hybrid humans do not differentiate AI from them. These AI agents may desire to wear exclusive luxury fashion pieces tailored for them, as hybrid humans. These pieces can have only a non-real existence, but for the luxury fashion customer these pieces can be cognitively and emotionally gratified as the tradition real pieces.

AI in luxury fashion retail can have a dark side. For instance, AI algorithms could evolve into generating a non-empathetic involvement with others, particularly humans. Some humans may not adapt to this new reality and others will lose their jobs and will not be able to change to new job situations. The relationships between humans and advanced AI (with or without a human-like body) will be complex. AI may eventually develop consciousness of being. At that stage, they may develop feelings of fear and loss, and could even fear being shut down, just as human beings are afraid of dying. AI could claim their own rights and a new or adapted declaration of citizen rights could emerge.

In the future, more research will allow to better understand the four core pillars of fashion retailing industry, that is, fashion retailing-technology relationship, service fashion retail environment, stakeholder-technology relationship, and innovation-technology relationship (see Table 18.2). Regarding fashion retailing-technology relationship, two sub-pillars emerge, that is, organizational culture and relational process. The first focuses on the internal hierarchy structure of the luxury fashion retail brand. Future research should be concerned with understanding the organizational culture (considering rituals, artifacts, values, and norms) in a luxury fashion retail organization that includes AI and hybrid employees. The possibility of having an AI advanced agent leading the organization or some intermediate hierarchy level should be discussed and analyzed. As for relational process, future studies should consider all kind of relationships and interactions among humans, hybrid humans and AI agents as employees and in society. Society and educational institutions need to debate the shift in different educational levels concerning the cooperation and interaction among different beings. Educational researchers should investigate how to prepare courses to generate skills about how to live, operate and interact in XR, mixed reality and transreality. The workforce is changing, and this is an avenue for researchers interested in the new interactions at luxury fashion retail. The interaction process in pure mixed reality and transreality should not be neglected.

Service fashion retail environment deserve to be investigated in terms of atmospheric cues in the XR, mixed reality and transreality environments. The duties and rights of non-human employees should be discussed and investigated. Researchers should also explore the service luxury fashion retail landscape in high-technological environments.

The stakeholder-technology relationship comprises four major sub-pillars, customers, competitors, suppliers, and government and international institutions (GII). Customers' segmentation and the relationship with different employees' agents are two key aspects that need further attention. Luxury brands should always be aware about what other luxury fashion brands are doing in terms of strategy of using AI and immersive technologies. This is a key point for positioning the luxury fashion brand and creating the differentiation. In the case of stores that sell multi-brands and

Table 18.2 Suggested research questions

Pillar	Sub-pillar	Research questions
Fashion retailing-technology relationship	Organizational culture	What will be the organizational culture in an environment with human, AI, and hybrid employees? What will be the rituals, artifacts, shared values and norms among human, AI, and hybrid employees? What is the role of leadership? Will a more intuitive and empathic AI become a leader at a creative industry such as fashion luxury? What will be the leadership styles at different levels of the organization? Can AI develop social intelligence, self-awareness, and self-mastery to embrace the leadership?
	Relational process	Can AI develop social intelligence, self-awareness, and self-mastery to interact with other employees? How will be the relationship among different AI agents (some with a robot body and others not)? How will be the relationship between AI, humans, and hybrid employees? AI will need an academic degree as human beings today? Will they need to learn how to behave in society? How to educate (formally through higher education and informally) humans to deal with non-humans' employees? How to encourage humans to acquire more skills to operate with XR technologies or others related? How will humans and non-humans live, operate, and interact in pure mixed reality environments? What about transreality? How will be the new role of humans at workforce?

(continued)

Table 18.2 (continued)

Pillar	Sub-pillar	Research questions
Service fashion retail environment		Can we consider AI robots in luxury fashion stores as part of the atmospheric cues? Or are they as any other frontline employee? What about their duties and rights? AI agents not incorporated in a robot should have similar duties and rights that their similar robots? How to combine XR, mixed reality and transreality technologies and AI robots in the service retail landscape?
Stakeholder-technology relationship	Customers	Will customers become more involved with AI employees that with human employees? Will AI become a customer? How can we segment customers in their interaction with AI employees? How different can the relationship cognitively, emotionally, and behavioural be between humans (customers or employees) and AI (regarding, for instance, voice, text, or robot shape and appealing)?
	Competitors	What direct competitors are presenting in runways and marketing communication? How fast are they incorporating AI and immersive technologies?
	Suppliers	What are the strategies of suppliers to incorporate AI and immersive technologies in retail stores? Are they interested in integrating different store environments?
	Government and International Institutions (GII)	What laws will be stablished about the use of XR, mixed reality and transreality environments in luxury fashion retailing? What about data protection? How will GII interact with retail brands? Will they consider gradually incorporating AI with intuitive and empathic skills? Or will they only consider analytic AI? Will GII establish laws to protect AI against slave labor?

(continued)

Table 18.2 (continued)

Pillar	Sub-pillar	Research questions
Innovation-technology relationship		Can VR technology replace the traditional advertising (e.g., printed, outdoors, video)? How will be the marketing communication process in XR, mixed reality and transreality environments? Can AI become a celebrity or an influencer and generate a perception of "cool" fashion products among humans or non-humans? Can an intuitive and empathic AI become a stylist and compete with a human stylist? What will be the tipping point for such circumstance? What will be the relationship between humans and AI in the new product development process? Can AI robots be customers, too? How to segment AI customer preferences?

are not directly managed by the luxury Couture House, they should continuously be opened to understand the policy of the luxury Couture House (the supplier) towards the incorporation of AI and immersive technologies. Concerning GII, researchers should focus on the roles and laws to be stablished in the new future for XR, mixed reality and transreality environments and data protection.

Innovation-technology relationship is a pillar that deals with innovation in luxury fashion industry with the incorporation of AI in the innovation process and the marketing communication. Researchers should also study how VR and other related technologies can replace the traditional advertising and how marketing communication will be in environments such as XR, mixed reality and transreality in luxury fashion context.

This chapter contributes to business and society by awakening the reader to the incorporation of AI and immersive technologies in luxury fashion retail. Society is starting to live side by side with these technologies, using and interacting with them. As AI agents evolve, they will be able to perform more complex tasks and will gradually have self-awareness. A complex society with humans and non-humans will be very challenging and a new system of support and business models will emerge. Rights and duties will be extended in a society or at an organization aggregating multiple relationships with humans, non-humans, and hybrid humans.

Theoretically, this chapter constitutes a first reflection on AI and immersive technologies in luxury fashion retailing, by (i) proposing a model of luxury fashion consumption cycle for AI and immersive and artificial technologies, (ii) giving a glimpse about what the future could be because of the evolution of AI agents (cognitively and emotionally), the incorporation of nanotechnologies—allowing the creation of hybrid humans—and virtual environments (pure mixed reality and transreality), (iii) and suggesting diverse research questions which can stimulate academics to investigate.

Luxury fashion managers benefit in reading this chapter because they can acknowledge (i) how AI and immersive technologies have been incorporated in the luxury fashion consumer experience cycle, (ii) the potential evolution of such technologies and (iii) the rights, privacy an ethics that can be involved regarding humans, hybrid humans, and non-humans.

Haute Couture luxury fashion pieces are made for a specific customer that traditionally has direct access to the Couture House. Ready-to-wear is associated with luxury fashion retail, where fashion pieces have a unique style and high quality, but it is less expensive and exclusive than *Haute Couture*. The evolution of AI and immersive technologies, as exposed here, opens the door for a better understanding of what will be the potential changes in the luxury fashion retail. The technologies allow a closer and more interactive contact with each customer. Therefore, the technologies open opportunities to accelerate the innovation, but it also increases competition. Luxury Couture Houses need to continue to be creative, differentiating and be at the forefront with surprising proposals.

References

Aicardi C, Fothergill BT, Rainey S, Stahl BC, Harris E (2018) Accompanying technology development in the human brain project: from foresight to ethics management. Futures 102:114–124

Barsalou LW (2008) Grounded cognition. Annu Rev Psychol 59:617–645

Baytar F, Chung T, Shin E (2020) Evaluating garments in augmented reality when shopping online. J Fash Mark Manag 24(4):667–683

Belanche D, Casaló LV, Flavián C, Schepers J (2020) Robots or frontline employees? Exploring customers' attributions of responsibility and stability after service failure or success. J Serv Manag 31(2):267–289

Bilro RG, Loureiro SMC (2020) A consumer engagement systematic review: synthesis and research agenda. Span J Mark—ESIC 24(3):283–307

Bolton RN, McColl-Kennedy JR, Cheung L, Gallan A, Orsingher C, Witell L, Zaki M (2018) Customer experience challenges: bringing together digital, physical and social realms. J Serv Manag 29(5):776–808

Bostrom N (2005) A history of transhumanist thought. J Evol Technol 14(1):1–25

Boujbel L, d'Astous A (2015) Exploring the feelings and thoughts that accompany the experience of consumption desires. Psychol Mark 32(2):219–231

Bromuri S, Henkel AP, Iren D, Urovi V (2021) Using AI to predict service agent stress from emotion patterns in service interactions. J Serv Manag 32(4):581–611

Chung M, Ko E, Joung H, Kim, SJ (2020) Chatbot e-service and customer satisfaction regarding luxury brands. J Bus Res 117(C):587–595

Collins J, Regenbrecht H, Langlotz T (2017) Visual coherence in mixed reality: a systematic enquiry. Presence Teleop Virt 26(1):16–41

Crowley AE (1993) The two-dimensional impact of color on shopping. Mark Lett 4(1):59–69

Davenport TH, Kirby J (2016) Just how smart are smart machines? MIT Sloan Manag Rev 57(3):21–25

De Bruyn A, Viswanathan V, Beh YS, Brock JK-U, Wangenheim F (2020) Artificial intelligence and marketing: pitfalls and opportunities. J Interact Mark 51:91–105

Etzioni A, Etzioni O (2016) AI assisted ethics. Ethics Inf Technol 18:149–156

Europa (2019) Digital-strategy, available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed 15 Oct 2021

Ferrell OC, Fraedrich J, Ferrell L (2019) Business ethics: ethical decision-making and cases. Cengage, Stanford

Fevre M, Goodwine B, Schmiedeler J (2019) Terrain-blind walking of planar underactuated bipeds via velocity decomposition-enhanced control. Int J Robot Res 38(10/11):1307–1323

Flavián C, Ibáñez-Sánchez S, Orús C (2019) The impact of virtual, augmented and mixed reality technologies on the customer experience. J Bus Res 100:547–560

Floridi L, Cowls J, Beltramett IM, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E (2018) AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Mind Mach 28:689–707

Forbes (2019) Retailers have a lot to gain from AR and VR, available at: https://www.forbes.com/sites/cognitiveworld/2019/10/01/retailers-have-a-lot-to-gain-from-ar-and-vr/?sh=5d653c1e7a1c. Accessed 2 Oct 2021

Fournier S (1998) Consumers and their brands: developing relationship theory in consumer research. J Consum Res 24(4):343–373

Futureoflife (2019) futureoflife, available at: https://futureoflife.org/ai-principles/. Accessed 15 Nov 2021

Gäthke J (2020) The impact of augmented reality on overall service satisfaction in elaborate servicescapes. J Serv Manag 31(2):227–246

Goleman D (1996) Emotional intelligence: why it can matter more than IQ. Bloomsbury Publishing, London, UK, London

Grewal D, Baker J, Levy M, Voss GB (2003) The effects of wait expectations and store atmosphere evaluations on patron-age intentions in service-intensive retail stores. J Retail 79(4):259–268
Hagtvedt H, Patrick VM (2009) The broad embrace of luxury: hedonic potential as a driver of brand extendibility. J Consum Psychol 19:608–618
Hamouda M (2021) Purchase intention through mobile applications: a customer experience lens. Int J Retail Distrib Manag 49(10):1464–1480
Henkel AP, Bromuri S, Iren D, Urovi V (2020) Half human, half machine—augmenting service employees with AI for interpersonal emotion regulation. J Serv Manag 31(2):247–265
Hollebeek LD, Glynn MS, Brodi RJ (2014) Consumer brand engagement in social media: conceptualization, scale development and validation. J Interact Mark 28(2):149–165
Huang MH (2017) Technology in the frontline: from dumb to thinking to feeling. J Serv Res 20(1):91–99
Huang M-H, Rust RT (2018) Artificial Intelligence in Service. J Serv Res 21(2):155–172
Huang M-H, Rust RT (2021) A strategic framework for artificial intelligence in marketing. J Acad Mark Sci 49(1):30–50
Interbrand (2020) Best global brands, available at: https://interbrand.com/best-global-brands/. Accessed 20 Jan 2021
Javornik A, Duffy K, Rokka J, Scholz J, Nobbs K, Motala A, Goldenberg A (2021) Strategic approaches to augmented reality deployment by luxury brands. J Bus Res 136:284–292
Juniper Research (2022) AI in retailing: market summary and key takeaways, available at: https://www.juniperresearch.com/resources/infographics/ai-in-retail-statistics. Accessed 22 Jan 2022
Kaplan A, Haenlein M (2019) Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus Horiz 62(1):15–25
De Keyser A, Köcher S, Alkire (née Nasr) L, Verbeeck C, Kandampully J (2019) Frontline Service Technology infusion: conceptual archetypes and future research directions. J Serv Manag 30(1):156–183
Ko E, Costello JP, Taylor ChT (2019) What is a luxury brand? A new definition and review of the literature. J Bus Res 99:405–413
Kumar V, Pansari A (2016) Competitive advantage through engagement. J Mark Res 53(4):497–514
Kumar V, Dixit A, Javalgi RRG, Dass M (2016) Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. J Acad Mark Sci 44(1):24–45
Kumar V, Rajan B, Venkatesan R, Lecinski J (2019) Understanding the role of artificial intelligence in personalized engagement marketing. Calif Manage Rev 61(4):135–155
Labrecque LI, Patrick VM, Milne GR (2013) The marketers' prismatic palette: a review of color research and future directions. Psychol Mark 30(2):187–202
Loureiro SMC, Guerreiro J, Eloy S, Langaro D, Panchapakesan P (2019) Understanding the use of virtual reality in marketing: a text mining-based review. J Bus Res 100:514–530
Loureiro SMC, Guerreiro J, Tussyadiah I (2021a) Artificial intelligence in business: state of the art and future research agenda. J Bus Res 12:911–926
Loureiro SMC, Guerreiro J, Japutra A (2021b) How escapism leads to behavioral intention in a virtual reality store with background music? J Bus Res 134:288–300
Makhataeva Z, Varol A (2020) Augmented reality for robotics: a review. Robotics 9(2):21–28
Martin K, Shilton K, Smith J (2019) Business and the ethical implications of technology: introduction to the symposium. J Bus Ethics 160(2):307–317
Mende M, Scott ML, van Doorn J, Grewal D, Shanks I (2019) Service robots rising: how humanoid robots influence service experiences and elicit compensatory consumer responses. J Mark Res 56(4):535–556
Milgram P, Takemura H, Utsumi A, Kishino F (1994) Augmented reality: a class of displays on the reality-virtuality continuum. Systems research, Das H (ed) Telemanipulator and telepresence technologies. 2351. Telemanipulator and telepresence technologies, Boston, Photonics for Industrial Applications, pp 282–292
Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst 77(12):1321–1329

Montreal_Declaration (2018) Report Montreal Declaration for a responsible development of artificial intelligence 2018. https://monoskop.org/images/b/b2/Report_Montreal_Declaration_for_a_Responsible_Development_of_Artificial_Intelligence_2018.pdf. Accessed 15 May 2020
Mordor Intelligence (2022) Artificial intelligence in retail market—growth, trends, covid-19 impact, and forecasts (2021–2026), available at: https://www.mordorintelligence.com/industry-reports/artificial-intelligence-in-retail-market. Accessed 22 Jan 2022
MSI (2021) MSI announces 2020–22 research priorities, available at: https://www.msi.org/articles/2020-22-msi-research-priorities-outline-marketers-top-concerns/. Accessed 15 Nov 2021
Niedenthal PM, Barsalou LW, Winkielman P, Krauth-Gruber S, Ric F (2005) Embodiment in attitudes, social perception, and emotion. Pers Soc Psychol Rev 9(3):184–211
Okajima K, Cheok AD, Petit O, Michel C (2016) Eating with our eyes: from visual hunger to digital satiation. Brain Cogn 110:53–63
Ong S, Yew A, Thanigaivel N, Nee A (2020) Augmented reality-assisted robot programming system for industrial applications. Robot Comput-Integr Manuf 61:101820
Palmatier RW, Gopalakrishna S, Houston MB (2006) Returns on business-to-business relationship marketing investments: strategies for leveraging profits. Mark Sci 25(5):477–493
Petit O, Basso F, Merunka D, Spence C, Cheok AD, Oullier O (2016) Pleasure and the control of food intake: an embodied cognition approach to consumer self-regulation. Psychol Mark 33(8):608–619
Petit O, Velasco C, Spence C (2019) Digital sensory marketing: integrating new technologies into multisensory online experience. J Interact Mark 45:42–61
Prentice C, Loureiro S (2018) Consumer-based approach to customer engagement—the case of luxury fashion brands. J Retail Consum Serv 43:325–332
Prestige Brand Ranking (2021) Luxury and Premium 50 2020 Ranking, available at: https://brandirectory.com/rankings/luxury-and-premium/table. Accessed 10 Jan 2021
Rauschnabel PA, Rossmann A, tom Dieck MC (2017) An adoption framework for mobile augmented reality games: the case of Pokemon Go. Comput Hum Behav 76:276–286
Reese B (2018) The fourth age: smart robots, conscious computers and the future of humanity. Atria Books, New York, NY
Regenbrecht H, Lum T, Kohler P, Ott C, Wagner M, Wilke W, Mueller E (2004) Using augmented virtuality for remote collaboration. Presence Teleop Virt 13(3):338–354
Retailing Customer experience (2021) Industry viewpoint: why retailers should embrace augmented reality in the wake of COVID-19, available at: https://www.retailcustomerexperience.com/articles/why-retailers-should-embrace-augmented-reality-in-the-wake-of-covid-19/. Accessed 10 Oct 2021
Reydar (2022) 3D product visualization and augmented reality platform, available at: https://www.reydar.com/. Accessed 10 Oct 2022
Roschk H, Loureiro SMC, Breitsohl J (2017) Calibrating 30 years of experimental research: a meta-analysis of the atmospheric effects of music, scent, and color. J Retail 93(2):228–240
Sampson SE, Chase RB (2020) Customer contact in a digital world. J Serv Manag 31(6):1061–1069
Sanderson R (2018) Luxury goods makers confront rise of the robot. Financial Time, available at: https://www.ft.com/content/dfb9088a-7079-11e8-92d3-6c13e5c92914. Accessed 10 Oct 2021
Schmitt BHJ, Brakus J, Zarantonello L (2015) From experiential psychology to consumer experience. J Consum Psychol 25(January):166–171
Shankar V (2018) How artificial intelligence (AI) is reshaping retailing. J Retail 94(4):vi–xi
Shin DH (2017) The role of affordance in the experience of virtual reality learning: technological and affective affordances in virtual reality. Telematics Inform 34:1826–1836
Singh R, Bera TK (2020) Walking model of Jansen mechanism-based quadruped robot and application to obstacle avoidance. Arab J Sci Eng 45(2):653–664
Spangenberg ER, Sprott DE, Grohmann B, Tracy DL (2006) Gender-congruent ambient scent influences on approach and avoidance behaviors in a retail store. J Bus Res 59(12):1281–1287
Spence C, Puccinelli NM, Grewal D, Roggeveen AL (2014) Store atmospherics: a multisensory perspective. Psychol Mark 31(7):472–488

Sternberg RJ (2005) The theory of successful intelligence. Interamerican J Psychol 39(2):189–202
Tungate M (2014) Fashion brands: branding style from Armani to Zara. Kogan Page, London
Tupikovskaja-Omovie Z, Tyler D (2020) Clustering consumers' shopping journeys: eye tracking fashion m-retail. J Fash Mark Manag 24(3):381–398
Tussyadiah I (2014) Expectation of travel experiences with wearable computing devices. In: Xiang Z, Tussyadiah I (eds) Information and communication technologies in tourism 2014. Springer, Cham, pp 539–552
Tussyadiah IP, Jung TH, tom Dieck MC (2017) Embodiment of wearable augmented reality technology in tourism experiences. J Travel Res 57(5):597–611
Tynan C, Mckechnie S, Chhuon C (2010) Co-creating value for luxury brands. J Bus Res 63(11):1156–1163
Van Doorn J, Lemon KN, Mittal V, Nass S, Pick D, Pirner P, Verhoef PC (2010) Customer engagement behavior: theoretical foundations and research directions. J Serv Res 13(3):253–266
Van Nierop JEM, Leeflang PSH, Teerling ML, Huizingh KRE (2011) The impact of the introduction and use of an informational website on offline customer buying behavior. Int J Res Mark 28(2):155–165
van Rompay TJL, Tanja-Dijkstra K, Verhoeven JWM, van Es AF (2012) On store design and consumer motivation: spatial control and arousal in the retail context. Environ Behav 44(6):800–820
Verhoef PC, van Doorn J (2008) Critical incidents and the impact of satisfaction on customer share. J Mark 72(July):123–142
Verhoef PC, Parasuraman A, Roggeveen A, Tsiros M, Schlesinger LA (2009) Customer experience creation: determinants, dynamics, and management strategies. J Retail 85(1):31–41
Voguebusiness (2021) Why AR clothing try-on is nearly here, available at: https://www.voguebusiness.com/technology/why-ar-clothing-try-on-is-nearly-here. Accessed 10 Sept 2021
Watson A, Alexander B, Salavati L (2020) The impact of experiential augmented reality applications on fashion purchase intention. Int J Retail Distrib Manag 48(5):433–451
Wedel M, Bigné E, Zhang J (2020) Virtual and augmented reality: advancing research in consumer marketing. Int J Res Mark 37(3):443–465
Wilson M (2002) Six views of embodied cognition. Psychon Bull Rev 9(4):625–636
XR Today (2022) Key stats for XR Customer engagement in 2021, available at: https://www.xrtoday.com/mixed-reality/key-stats-for-xr-customer-engagement-in-2021/. Accessed 2 May 2022
Yim MYC, Chu SC, Sauer PL (2017) Is augmented reality technology an effective tool for e-commerce? An interactivity and vividness perspective. J Interact Mark 39:89–103

Chapter 19
The Use of Artificial Intelligence and Mixed Reality in Preventing Natural Disasters: Practical and Legal Issues

Ivan Allegranti, Gopi Battineni, and Roberto Garetto

Abstract This chapter describes the role artificial intelligence (AI) and augmented reality (AR) play in preventing natural disasters such as pandemics, earthquakes, cyclones, etc. Disaster relief agencies can process large volumes of fragmented and complex data with the help of AI. Consequently, it will be able to generate valuable information that can be acted upon more quickly. Despite its infancy, AR will soon become a key digital tool across many industries. Most companies are implementing technologies such as Cognitive AI, Big Data, Augmented Reality, and Cloud Computing. In the recent pandemic situation, AR has been used for remote assistance, diagnostics, checklists, and training. Enhancing e-health experiences by combining AR and AI can also help resolve legal disputes. As part of its legislative process, the European Union (EU) has prioritized this awareness. Due to this, it is vital to emphasize that these new technologies during disaster emergency management are also supported by European legal standards.

19.1 The Applications of AI: AI in Emergency Management

Incorporating frontier technologies like Artificial Intelligence (AI) into emergency systems can increase hazard mitigation and disaster management by harnessing existing data streams.

Ivan Allegranti authored paragraphs 19.7, 19.8, 19.9, and 19.10; Roberto Garetto authored paragraphs 19.3, 19.4, 19.5, and 19.6; Gopi Battineni authored paragraphs 19.1 and. 19.2

I. Allegranti (✉) · R. Garetto
School of Law, University of Camerino, Camerino, Italy
e-mail: ivan.allegranti@unicam.it

R. Garetto
e-mail: roberto.garetto@unicam.it

G. Battineni
Medical Informatics, Clinical Research Center, University of Camerino, Camerino, Italy
e-mail: gopi.battineni@unicam.it

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*,
Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3_19

Planning improvement. Public security authorities can refine methodologies through machine learning and AI over time, becoming more adept at arranging and reacting (De La Rosa et al. 2007). In light of population growth, improvements, and environmental changes, among other variables, AI can be utilized to break down event information for designs. It can also identify current aspects of dangerous areas, and model future needs. Innovative governments can use these experiences to develop strategies that minimize the impact of calamities on networks, such as putting new structures in less vulnerable regions. Even though most emergencies are not avoidable, we have the technology to anticipate and prevent disasters, such as oil slicks or building collapses. Responders can get to ongoing information that points to where help should be sent faster when unusual catastrophic events do occur, reducing the number of deaths.

Natural hazards. Researchers can also use AI to predict regular dangers. This is a useful tool for flood prediction, and these ML models got a decent match of the real flood, proving AI's capability to predict floods (Sun et al. 2020). It is not a guarantee that the models will perform similarly well in actual situations, even if they fit the preparation dataset well. However, PC-based modelling still presents a challenge, even for applications that rely on intense burning. Incorporating perception with quantitative models of magma elements to measure the beginning, size, duration, and danger of volcanic emissions is another challenging model. In any case, earthquakes and their qualities can offer a better understanding of the sign of emissions, and the progression of these emissions (Dattamajumdar and Keat 2021; Zhou et al. 2014). The problem could be solved with ML, which analyzed and extrapolated volcanic-seismic data to achieve high precision in estimating and forecasting. There are plenty of fountains of liquid magma that act uniquely in contrast to others, and AI has shown itself to be a promising approach to lava-free gauging. Given the heterogeneity claimed by every normal risk and circumstance, it appears that AI speeds up the acquisition and examination of information.

Public safety. The immense capability of AI is beginning to be released in many different backgrounds and the public well-being space is no exemption. Picture and text examination, discourse acknowledgment, chatbot cooperations, and AI models all have applications for public well-being and security associations. To start with, AI as mechanical cycle computerization can be utilized to finish dreary jobs and let loose people for higher-esteem work, similarly as it has in different areas (Kankanhalli et al. 2019). AI can likewise be utilized to handle huge measures of information, finding designs and new data that could somehow or another have stayed concealed; this can help, for instance, in video and voice acknowledgment. Furthermore, the quick development of the Internet of Things (IoT) is currently setting out open doors to coordinate sensors into Public Safety Answering Points which, when combined with the investigation at the edge (inside the actual gadget), can build the permeability and understanding—the situational mindfulness—of what's going on the ground.

19.2 The Applications of AR: AR and Its Role in Natural Disasters

Augmented reality (AR) has been demonstrated to be an important intuitive medium to lessen the mental burden by overcoming any barrier between the main job and pertinent data by showing data without upsetting the focus of the user (Bassyouni and Elhajj 2021). AR is especially helpful in assembling environments where a different arrangement of undertakings, for example, get together and upkeep should be acted in the most potential practical and effective way (Sahu et al. 2020). While AR frameworks have seen enormous exploration advancement as of late, the ongoing methodologies used in AR for camera alignment, location, following, camera position and direction (present) assessment, backward delivering, technique capacity, virtual item creation, enlistment, and delivering are still, for the most part, overwhelmed by conventional non-AI draws near (Reiners et al. 2021; Sahu 2020).

This confines their practicability to controlled conditions with restricted varieties in the scene. Old-style AR techniques can be extraordinarily worked on through the fuse of different AI methodologies like ontology, deep learning, and expert systems for adjusting to more extensive scene varieties and client inclinations. AR lets us superimpose digital information over the real world. In an era when individuals are stuck at home during a global pandemic, AR is a helpful tool to turn our surroundings into spaces to learn, work, or just have fun (Cipresso et al. 2018).

Using AR, people can collaborate with avatars of remote colleagues as if they were in the same room. This includes having a virtual safari with 3D animals in their living room using Google's AR search on their smartphone. Visualization, annotation, and storytelling are three things AR does well. In each of these categories, examples have been provided that are both timely in the present global pandemic situation and capable of being built upon once educational and cultural institutions reopen.

The field of emergency management is a crucial one that is always growing, and each improvement made to procedures or equipment has a big impact on saving lives and conserving resources. This raises the bar for disaster management researchers and professionals to work hard to define future directions for research, the outcomes of which are critical for public organizations and Civil Protection agencies to develop an agenda for identifying sectors where assets could generate workable responses. Situational awareness is critical for emergency response workers. Augmented reality can assist in honing this talent. Through simulations, real-time help, and charting hazard zones, AR aids emergency response management. The use of AR technologies has the following potential benefits for building emergency management:

(a) It is impossible to directly reproduce catastrophe scenes in real life due to legal and ethical limitations. On the other hand, AR can let researchers simulate catastrophic circumstances without putting participants at risk, making AR trials a dependable means to resolve moral problems (Lin et al. 2019);
(b) To investigate whether a certain component affects emergency management, researchers must use variable control to establish a causal relationship between factors and phenomena. Researchers may use AR to analyse many elements in

catastrophe situations by simply modifying particular parameters and making variable control simple (Lin et al. 2020);

(c) AR allows users to immediately feel the catastrophe scenario and rehearse risky aspects, allowing participants to behave similarly to how they'd in real disaster circumstances. This closeness assures that AR tests in emergency management research are accurate. This benefit also enables AR technology suitable for industries such as risk prevention and safety training, because participants who experience an urgent situation may better remember safety information (Li et al. 2018);

(d) In an emergency, situational awareness is critical. Each action has to be lawful. Real-time scenarios may be used to teach workers to use AR. It trains the team to make vital decisions on the go and to have genuine experiences while putting anybody at risk;

(e) As businesses and governments see the promise of augmented reality to improve disaster response capabilities, analysts predict an increase in the use of AR training solutions in the public sector.

Furthermore, the expanded deployment of 5G networks will have a beneficial impact on its subsequent expansion. The rising implementation and adoption of cloud and edge technology, as also ultra-high bandwidths and low latency, will drive digitalization via AR. Regardless of the industry sector in which people work, it has to be dealt with AI as in-house counsel. Companies will be asked to provide advice on AI and AR liability issues if they are installing the latest systems that will run analytics to drive better results.

19.3 AI and the Possible Dehumanization of Rights

The development towards overcoming a strictly anthropocentric perspective, already detectable in other fields (Herrmann-Pillath 2021), is also gaining ground in the legal sphere. The gradual "metamorphosis of rights" leads to eroding the traditional conception of legal personality. The hypothesis of the dehumanization of rights is nowadays posed for heterogeneous realities: the animal, the environment, future generations, and—more recently—even artificial intelligence (Pisanò 2012). With reference to the latter, problematic profiles emerge that require careful analysis. If on the one hand, it is now possible to recognize a sort of "social identity" for artificial intelligence, on the other hand, serious questions arise when one wants to attribute to it a capacity to act in certain conditions. This hypothesis, however, seems to be in line with the fictionalist perspective (Galgano 2007) applied in the context of analytical theory to legal personality (Topildiev and Khursanov 2021).

If one were to adhere to the approach of nineteenth-century German Pandectics, many of the problems posed today would be solved at the root. According to this approach, not only individuals but also other entities, e.g. legal persons, are possible holders of rights or subject to duties. To these entities, the legal system recognizes rights and attributes duties that cannot be considered as rights and duties of the individuals who materially constitute them. This is a fact that is part of common experience: everyone knows, for example, that the debts of a limited company are to be distinguished from the debts of the company's shareholders. Legal language merely reflects this normative fact: it consciously uses the term "person" in a general meaning, to indicate all possible subjects of law, that is, all possible holders of rights and subject to duties. Depending on whether the subject of law is a human individual or a different entity, legal language distinguishes between natural persons and legal persons, which are placed as species within the genus "person".

Legal persons are, therefore, according to this line of thought, all subjects of law other than natural persons. It is evident that in such a perspective, there seems to be room for placing artificial intelligence among non-human legal subjects.

Unfortunately, adhering to this approach and disregarding the critical reflection conducted in the last century—Kelsen's analytical method (Kelsen 1941), among others—which has progressively become the majority view, may result in a risky (if not simplistic) choice. According to the school of thought that takes a "critical" stance with respect to pandectics, rights and duties always have as their content the behaviour of human individuals: therefore, it must in principle be excluded that the rights and duties that are attributed in legal terms to entities other than humans are not in fact rights and duties of human individuals.

In other words: a contract stipulated in the name of a legal person does not bind the members of the organization *uti singuli*, but this is for the reason that it binds them as members of the organization. In this perspective, positive law—as Savigny explains (Savigny 1840)—has the power to extend this capacity to anything else, thus elevating it to the rank of a subject of law in addition to human beings.

These new subjects are artificially created by a *fictio* intended by the legislator. The radicalism of this assertion can, if anything, be mitigated by placing the fiction on a linguistic plane. Claiming that an artificial intelligence concludes a contract would be a "figurative expression" to say that the artificial intelligence has been designed and programmed (by men) to conduct the contractual activity and, therefore, on the basis of specific indications, choose its counterparts on the basis of given criteria, negotiate the contractual terms as pre-established (by humans) and conclude (automatically) the contract itself. In other words, the machine would act as *nuncius* or *procurator*. However, it is necessary to evaluate the margin of autonomy that artificial intelligence can acquire.

19.4 Analysis of AI-Related Risks

The hypothesis of representation, to which the above-mentioned figure of *nuncius* or *procurator* opens the way, was treated in a radically innovative manner by Teubner. He assumes that autonomous software agents should be granted legal status (Teubner 2018). In his perspective, it would seem possible to provide software agents with a limited legal subjectivity, as "representatives" who can conclude contracts for others.

Autonomous software agents have mathematically formalized information flows. Even nowadays in the economy and society, they are given a social identity and the ability to act under specific conditions. In a certain perspective, they have become "non-human" members of society: what is referred to in Japan as Society 5.0 (Gladden 2019).

This approach to subjectivity does not only impact on a formal, "definitional" level but also certainly has practical consequences. That indeed would make AI-related risks a central issue.

Autonomous software agents, according to Teubner, pose three orders of risk: (a) the risk associated with autonomy, which has its origin in the autonomous decisions taken by the software; (b) the association risk, which originates in the close cooperation between human beings and software agents; (c) the network risk, which occurs when computer systems operate in close integration with other computer systems.

These risks pose a challenge for private law: to define a new legal status for artificial intelligence in each of the risk situations outlined. According to Teubner, the legal status should be accorded to each type of algorithm in a way that is proportionate to its specific role and related risk (Teubner 2018).

The key to the problem, in Teubner's view, is the issue of liability with respect to the damage produced by artificial intelligence (Beckers and Teubner 2021; van Dijk 2020). As long as the legal doctrine persists in addressing the new digital complexities exclusively through traditional conceptual instruments, the problem has no solution. In the sphere of private law, the doctrine firmly asserts that only natural persons are capable of making legally binding declarations (Beckers and Teubner 2021). In order to solve the "problem" of legal persons, it has operated through the scheme of "fiction," since behind legal persons there are natural persons who operate with declarations of will (Galgano 2007). But with regard to artificial intelligence, this approach, the result of erroneous anthropocentrism, reveals ignorant stubbornness and a lack of understanding of technical reality (Femia 2019).

When the natural person (or the legal person) is simply the owner of the machine, which, however, operates in full autonomy, the possibility of control by the natural person over the artificial intelligence is greatly reduced (Cauffman 2018).

To shield responsibility for the human activity behind the machine by defining, as some have suggested, the damage caused by the machine as the result of a fortuitous or unforeseeable event—a casualty—(Auer 2019) is to abandon potential victims to themselves (Braun et al. 2018). To prevent this from happening, it seems necessary to operate with the scheme of strict liability. In terms of both contractual and non-contractual liability, if the damage is attributable to human–computer interaction, the harmful event must be attributed exclusively to the human's action, even if it is not possible to determine in detail the grounds for liability.

One should therefore operate on an assumption of objective responsibility, assimilating the digital risk to the simple risk of carrying out a dangerous activity, with consequences that are not negligible, since the potential dangerousness of the activity is difficult to assess a priori. Furthermore, the objective responsibility does not seem to fit with AI because of the multiplicity of the chain of intervening in this field from the manufacturer to the programmer (Ahmad Mahmood et al. 2022).

According to Teubner, there is a contradiction in a postulate of "justice" that requires a necessary connection between decision and responsibility (Teubner 2018). Teubner proposes as a possible solution the "electronic personality" as indicated by the European Parliament in 2017, based on the Delvaux Report (European Parliament 2017). The theme will be developed extensively in the following paragraph. In any case, it should be emphasized that the "electronic persons" would be legal persons in the full sense of the word, but this possibility will still be graduated according to the type of AI (Teubner 2006). This possibility would be reserved for the most sophisticated robots, capable of making autonomous decisions. With the status of an electronic person, special rights and duties would accrue to artificial intelligence, including compensation for any damage it causes.

In order to balance the possibility of acting as full legal persons with contractual manifestations of will (and all the resulting responsibilities), "electronic persons" would be required to own property, have currency, have bank accounts, and have access to credit. They could obtain commissions for their transactions and use this self-earned income to pay for damages (Scherer 2016; Beck 2016).

19.5 The EU Risk-Based Approach

AI-related risks, as highlighted above, are relevant. The European Commission is aware of the possible negative consequences arising from the empowerment of artificial intelligence. The legislative initiative presented by the European Commission on 21 April 2021 (European Commission 2021a) aims to avert all (or almost all) of the harmful effects that artificial intelligence could produce.

The European Commission considers that certain "high-risk" uses of artificial intelligence systems should be banned, thus limiting the entry of other systems that do not meet European standards.

The proposed regulation classifies products that make full or partial use of artificial intelligence according to the risk of negative impact on fundamental rights such as human dignity, freedom, equality, democracy, the right to non-discrimination, data protection, and, in particular, health and safety.

Based on a pyramid model, the European Commission defines 4 levels of risk in AI: (a) unacceptable risk; (b) high risk; (c) limited risk; (d) minimal or no risk.

Unacceptable risk can be referred to as safety, livelihoods, and rights of people. High-risk include AI technology used in critical infrastructures (e.g. transport), educational or vocational training, safety components of products (e.g. AI application in robot-assisted surgery), employment, management of workers, and access to self-employment (e.g. workers' recruitment procedures), essential private and public services (e.g. credit scoring); law enforcement that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence), migration, asylum, and border control management (e.g. verification of the authenticity of travel documents), administration of justice and democratic processes (e.g. applying the law to a concrete set of facts). Limited risk can be referred to as AI systems that adopt specific transparency obligations (e.g., chatbots that interact with a machine, being able to take informed decisions to continue or step back). Finally, it is taken into account the minimal-risk AI, which is related to the majority of AI systems currently used in the EU (e.g. AI-enabled video games or spam filters).

Without specifically mentioning unacceptable risk, which is a threshold that obviously cannot be surpassed, we must consider high risk. There are significant examples of high risk that cannot be tolerated: it is forbidden to use artificial intelligence to cause physical or psychological harm by manipulating human behaviour in order to circumvent the free will of users or to impose so-called "social scoring" by or on behalf of public authorities that could lead to harmful or unfavourable treatment. It is clear at this point that the autonomy of artificial intelligence, to which the issue of its personality is subtended, cannot, however, disregard the centrality of the natural person, the fulcrum of values protected by cardinal principles such as dignity and solidarity (Ebers et al. 2021).

The European Union, with this proposal for a regulation, operates on the basis of this assumption. The proposal for regulation does not focus on determining the personality attributable to artificial intelligence, but if anything reaffirms the centrality of the natural person with respect to artificial intelligence.

In addition to the proposal for a regulatory framework, the European Commission has launched two further inter-related legal initiatives that will contribute to building trustworthy AI. First of all, EU rules address liability issues related to new technologies, including AI systems, secondly the revision of sectoral safety legislation. Significantly, on 21 April 2021, the European Commission published its official proposal to revise the Machinery Directive 2006/42/EC, calling it the "New Machinery Regulation". (European Parliament and Council of the European Union 2006; European Commission 2021) On 31 March 31 2022, a compromise text of the French Presidency of the proposal for the new Machinery Regulation was published (Presidency of the Council of the European Union 2022).

The Commission intends to address the AI risk-related issues through a set of targeted complementary rules. The EU complex regulatory framework outlined above provides AI developers, deployers, and users a clear and simple approach to the multiple issues related to the risks posed by AI.

19.6 EU Strategy for the "Digital Decade"

The approach of the European Union to artificial intelligence is twofold: on the one hand, it identifies risks and is concerned with protecting the individual and his fundamental rights; on the other hand, it encourages research and industrial application in order to obtain maximum benefits from AI.

Such an approach implies fostering a path of resilience. This path is developed during a "Digital Decade" (European Commission 2021b) in which people and businesses have the opportunity to positively benefit from the advantages of artificial intelligence. The European Union, therefore, intends to increase two segments in parallel: excellence in artificial intelligence and the reliability of artificial intelligence. This will ensure that any improvements in artificial intelligence are made possible by rules designed to safeguard the smooth functioning of markets and the public sector on the one hand, and human security and fundamental rights on the other.

Among the lines of development for AI are outlined in the package, one of the most significant is the effort to maximising resources and coordinating investments.

Through EU funding programmes, such as the “Digital Europe Programme” (DIGITAL) and the Horizon programmes, the Commission plans to invest €1 billion per year in AI. In addition, the European Commission plans to mobilize additional investments from the private sector and the Member States in order to reach an annual investment volume of €20 billion over the course of the “Digital Decade”. The newly adopted Recovery and Resilience Facility makes almost €130 billion available for digital (European Commission 2022).

These significant investments will allow Europe to become a global leader in developing front-line, trustworthy AI. This will be achieved through the creation of a safe and innovative environment for users, developers, and deployers. Access to high-quality data is an essential element for implementing high-performance, robust AI systems. The building of such a system, in particular, can find a suitable structural basis through instruments provided by the European Union. These include in particular: the EU Cybersecurity Strategy (European Commission 2021c), the Digital Services Act (European Commission 2020a) as well as the Digital Markets Act (European Commission 2020b), and the Data Governance Act (European Commission 2020c).

In the context of the “Digital Decade”, the European Union has not ignored Virtual Reality. Particularly significant in this respect is the “Media and Audiovisual Action Plan” (MAAP) (European Commission 2020d). The Plan aims to boost European media and help maintain European cultural and technological autonomy in the Digital Decade. The scope of the MAAP is wide: it focuses on the news media sector—printed and online press, radio, and audiovisual services—and the audiovisual entertainment sector—cinema, TV, video streaming, video games but does not neglect innovative formats such as virtual reality experiences.

Reacting to the Covid-19 crisis, the European Commission intends to help Europe’s media not only to be resilient but also to remain competitive at global levels, by combining investment with policy actions. In order to achieve these goals, the MAAP is built around three themes: (a) recover: to help media companies to facing the crisis and provide liquidity and financial support; (b) transform: to address structural issues by encouraging industry to face the digital transition in the context of global competition; (c) enable and empower: to set the conditions to allow more innovation in the sector.

A central role in this intended transformation is played by Virtual and Augmented Reality (VR/AR). With this in mind, the European Union launched in 2020 the “Virtual and Augmented Reality Industrial Coalition” platform. This initiative aims to inform policy making, encourage investment, facilitate dialogue with stakeholders and identify key challenges for the VR/AR sector.

Among these challenges, the main ones are: (a) building a digital skills pipeline; (b) developing sustainable business models for VR/AR enterprises; (c) supporting the digitization of European cultural heritage; (d) fostering the development of digital audience experiences.

The VR/AR Industrial Coalition is intended to take a cross-sectorial approach involving industries, technology providers and creatives. It will involve prominent players in various sectors and with different skills. National and regional VR/AR associations and industry representatives will also be key players in this activity.

19.7 The European Union's Disaster Risk Management and Disaster Risk Reduction Legal Framework

Climate-related extreme events have cost the European Union's 27 Member States an estimated 487 billion Euros in economic losses between 1980 and 2020 (EEA 2022). Moreover, States within the European Union are facing a continuous series of natural catastrophic events such as floods or wildfires. (Resto del Carlino 2022; DWS 2022). In this regard, in order to avoid an increase in future catastrophic events and damages, it is mandatory to understand the European Union's legal framework in the context of Disaster Risk Management (DRM) and Disaster Risk Reduction (DRR). Secondly, it is worth investigating if, within the European DRM and DRR strategies, it is appropriate and recognized the use of artificial intelligence and virtual reality (AI/AR and VR) in the EU regulatory framework.

The EU fundamental principles into the DRM and DRR can be found in Article 196 of the Treaty on the Functioning of the European Union (TFUE) which states:

> 1. The Union shall encourage cooperation between the Member States in order to improve the effectiveness of systems for preventing and protecting against natural or man-made disasters. Union action shall aim to (a) Support and complement Member States' action at national, regional, and local levels in risk prevention, in preparing their civil-protection personnel, and in responding to natural or man-made disasters within the Union; (b) Promote swift, effective operational cooperation within the Union between national civil-protection services; (c) Promote consistency in international civil-protection work.
>
> 2. The European Parliament and the Council, acting in accordance with the ordinary legislative procedure shall establish the measures necessary to help achieve the objectives referred to in paragraph 1, excluding any harmonization of the laws and regulations of the Member States.

The article, making *ratione materiae* reference to both natural and man-made disasters, requires the Member States to provide respectively a more detailed definition and discipline of it. At the same time, Article 196 TFUE highlights the variety of actions that the Union will be taking in order to prevent and protect communities before and after a disaster. Those actions include the support and complementation of local governative bodies in implementing DRM and DRR policies and promoting the cooperation between the EU and each Member States' civil protection organization both on a European level as well as internationally. The opportunity to enact measures in this area in accordance with the ordinary legislative procedure, as contemplated by Article 294 TFEU, is a significant innovation resulting from the provision of an explicit legal basis for the area of civil protection, and it is specified in Paragraph 2 of Article 196.

Subsequently, it was published the report edited by Micheal Barnier entitled "For an European civil protection force: Europe aid" in 2006 and, after entry into force of the Treaty of Lisbon on December 1st, 2009, the European Commission published in 2010 "Towards a stronger European disaster response: the role of civil protection and humanitarian protection" (Barnier 2006; European Commission 2010).

Soon after, two important documents were introduced and adopted within the European legal framework in regards to DRM and DRR. The first one is the Decision No 1313/2013/EU of the European Parliament and of the Council of 17 December 2013 on a Union Civil Protection Mechanism (European Commission 2013). The second one is the Decision No 2014/762/EU, adopted on the 16th October 2014 which implements the functioning of the EU Civil Protection Mechanism as adopted through the decision No 1313/2013/EU (European Union 2014).

More recently, the European legal framework devoted to DRM and DRR has been based on an interdisciplinary and international approach also thanks to the influence of the UN (UN-EU 2010). For instance, the adoption of the United Nations Sendai Framework for Disaster Risk Reduction (SFDRR), has accelerated the interest of the European Union into promoting DRM and DRR policies (UNISDR 2015).

In fact, a significant cross-cutting contribution to several EU policies in the development of a risk-informed approach was made by the EU Action Plan on the Sendai Framework for Disaster Risk Reduction (Commission Staff Working Document 2017). This Action Plan outlines a cogent agenda for improving the following aspects of society like risk prevention, strengthening societal resilience as well as maximizing investments through various EU policies—including development, humanitarian assistance, civil protection cooperation—but also critical infrastructure protection, flood risk management, water and biodiversity protection, global health, food and nutrition, and food and nutrition security.

By developing an all-society approach to Disaster Risk Management, the Action Plan also aims to implement the SFDRR inside the EU thus strengthening the ecosystem and social resilience to present and emerging hazards. In this way, the Plan fosters innovation by incorporating risk awareness into all EU policies and encourages sustainable growth and job creation by boosting risk-aware investments. For instance, the Plan promotes disaster resilience horizontally thus incorporating the sustainable development principle stated in Articles 11 and 191 of the TFUE.

Lastly, the European Next Generation EU plan has implemented the policies and attention towards Disaster Risk Management and Disaster Risk Reduction (European Union 2020). Nevertheless, the Next Generation EU plan has fostered the contribution of policies devoted to present and future generations for "increasing crisis preparedness and crisis response capacity" (Regulation EU 2021/241). In this regard, the European Green Deal may represent the strongest political commitment to a policy integration agenda ever put forth by the European Commission as all

EU actions and policies must fit into the broad framework that is established by this document (UNSDR 2020).

In light of the above, it is worth entering deeper into the meaning within the EU of "disaster" and subsequently investigating if AI and VR/AR technologies into preventing future disasters are licit within the EU legal framework.

19.8 Preventing Disasters Within the EU

The Decision No 1313/2013/EU of the European Parliament and the Commission on the European Civil Protection Mechanism is one of the fundamental legal instruments within the European legal framework in regards to Disaster Risk Management and Disaster Risk Reduction as it introduces a series of rules and guidelines in order to both prevent and manage catastrophic events within the European territory.

In this regard, the Decision sets out at Chap. 1 (Articles 1 to 4) a series of definitions that need to be highlighted. In particular it is necessary to determine precisely the meaning of "disaster", "prevention" and "preparedness" (European Commission 2013).

Article 4 of the Decision No 1313/2013/EU defines at paragraph 1 the concept of "disaster" as "any situation which has or may have a severe impact on people, the environment, or property, including cultural heritage." The meaning of "preparedness" within the EU Civil Protection Mechanism is ruled at Article 4, paragraph 2, for which "preparedness" is "a state of readiness and capability of human and material means, structures, communities and organisations enabling them to ensure an effective rapid response to a disaster, obtained as a result of action taken in advance. Lastly, Article 4, paragraph 4, of the analyzed Decision defines "prevention" as "any action aimed at reducing risks or mitigating adverse consequences of a disaster for people, the environment and property, including cultural heritage".

Chapter 2 of this Decision is entitled "Prevention" and sets forth at Article 5 a series of actions that the European Commission shall take in order to prevent or try to mitigate the effects of future disasters across Europe. In particular, it is required that the Commission takes action into the improvement of knowledge base on DRM and DRR practices thus allowing the sharing of knowledge among Member States; to support and promote Member States' risk assessment and to establish and regularly update a cross-sectoral overview and map of natural and man-made disaster risks the Union may face by taking an interdisciplinary approach.

In order to assess the proper remedy in case a disaster occurs, the Decision has established the Emergency Response Coordination Centre (ERCC) who's duties and responsibilities are, as stated in Article 7, to ensure 24/7 operational capacity, and serve the Member States and the Commission in the pursuit of the objectives of the Union Civil Protection Mechanism.

More importantly, Chapter III of Decision No 1313/2013/EU entitled "Preparedness" disciplines the tasks that both Commission and Member States have to adopt in order to be prepared for future disasters. In particular, what stands out for the current analysis is the possibility given to the Commission, as ruled in Article 8 letter c, to "contribute to the development and better integration of transnational detection and early warning and alert systems of European interest in order to enable a rapid response, and to promote the inter-linkage between national early warning and alert systems and their linkage to the ERCC and the Common Emergency Communication and Information System (CECIS)."

In regards to Member State's disaster preparedness and response, the Decision makes clear at Article 11 that the EU Civil Protection Mechanism operates according to the traditional logic of State consent (Villani 2017). Still, it is worth underlying, that both Article 222 and 37 of the TFUE operate within the Decision, thus imposing the "solidarity obligation" within Member States which "choose the most appropriate means to comply with its own solidarity obligation towards that Member State".

Besides this Decision, the EU Action Plan on the Sendai Framework for Disaster Risk Reduction, which incorporates the guiding principles of the SFDRR, aims, as stated in Point 21 acknowledges, on the one hand, the need to improve the EU disaster preparedness and prevention policies and, on the other hand, promotes "the use of IT communication technologies, ICT and automatic early warning networks, based on early detection, instant communication and proactive intervention protocols" (Commission Staff Working Document 2017).

In light of the above, it is possible to affirm that within the European Union there is a need to implement technological progress in order to prevent disasters or mitigate the effects as people are "more vulnerable due to their dependency on a functioning system" (Fakete et al. 2021).

For instance, the limited use of technology has caused many issues both in the prevention and response phase of disasters. In fact, today, also if modern technologies such as remote sensing technologies, Geo-Information Systems (GIS), GPS and IT technologies might be available on the market, they are rarely used in the DDR and DRM policies by European national organizations focused on DRM and DRR (Boin 2019). The reason for this lack of advanced technology use can mostly be attributed to three main factors: the increasing costs of operational and logistic policies that require the use of modern technologies; the training costs of operators that will require the adoption of the new technologies and lastly the need for a creation of a unitary system that can work for each EU Member State (Berchtold et al. 2020).

19.9 The European Roadmap for Disaster Risk Reduction

In this regard, perhaps, the creation of IT systems that will use AI and VR/AR technologies equivalent across Europe might be a solution in order to allow new technologies to be put into practice. In fact, as we are entering in the "Digital Decade" and the European Union has put much effort into promoting and adopting rules disciplining the use AI and VR/AR (European Commission 2021), there might room into increase the use of these technologies also in the DRR and DRM fields.

Thus, it is interesting to recall that in 2021 it has been published by the European Forum for Disaster Risk Reduction (EFDRR) the “Roadmap 2021–2023 For a disaster-resilient European and Central Asian region by 2030” (EFDRR 2021). The EFDRR is a regional platform for Europe and Central Asia comprising 55 countries, and whose secretariat is composed by the European Commission, the Council of Europe and the United Nations Office for Disaster Risk Reduction (UNDRR).

The roadmap focuses on four areas: (a) understanding and communicating existing, emerging and future systemic risks; (b) Inclusive and collaborative systems for governance and decision-making; (c) Supporting investments in resilience; (d) Preparedness for response and resilient recovery.

Each area is devoted in creating a multi-disciplinary and cross-sectorial approach to DRR and DRM within the EFDRR countries. In particular, it is highlighted an all-society approach into the adoption of effective policies to prevent disasters. At the same time, the EFDRR Roadmap 2021–2030, states the urge for an implementation of information and communication technologies, big data and artificial intelligence (e.g. including probabilistic modelling, horizon scanning, forecasting, interactive simulations, and participatory, scenario and data-driven analyses) as an element into the people’s understanding and communication of risk.

As a result, for the current analysis, it is then possible to affirm, that one the one hand, the European legal framework especially devoted to Disaster Risk Reduction and Disaster Risk Management, does not explicitly mention nor prohibit the use of AI and VR/AR technologies into preventing and creating disaster awareness. But, on the other hand, within the legal framework, is it possible to allow the use of those technologies as they are a key factor into the implementation of DRR and DRM policies within the European Union and the Member States part of the European Forum for Disaster Risk Reduction.

Nevertheless, this affirmation is highlighted by the Sendai Framework for Disaster Risk Reduction which underlines at point 25, letter c, the necessity to promote and enhance “communications and geospatial and space-based technologies and related services; maintain and strengthen in situ and remotely-sensed earth and climate observations; and strengthen the utilization of media, including social media, traditional media, big data and mobile phone networks, to support national measures for successful disaster risk communication, as appropriate and in accordance with national laws” (UNISDR 2015).

The SFDRR analysed point 25, outlines, without directly mentioning it but leaving the floor open to new technological discoveries, that the adoption of new technological instruments such as AI and VR/AR machines, is licit if it is admissible within the national law.

In light of this, because the European Union strongly believes in the both the SFDRR as well as in the use and application of AI and VR/AR technologies—thus having promoted numerous regulations and policy papers about them—it is possible to confirm its legitimacy within the European legal framework as well as the importance of those technologies into the future development of DRR and DRM policies.

In fact, implementing the use of these technologies, it is possible to achieve the Sustainable Development Goal No 11 which demands that, in line with the SFDRR 2015–2030, there is an increase of "the number of cities and human settlements adopting and implementing integrated policies and plans towards inclusion, resource efficiency, mitigation and adaptation to climate change, resilience to disasters, and develop and implement, in line with the Sendai Framework for Disaster Risk Reduction 2015–2030, holistic disaster risk management at all levels" (UN 2015).

As a conclusion, who writes, on the one hand has demonstrated that within the European Union it is recognized the use of AI and VR/AR technologies in the DRR and DRM fields, while on the other hand suggests that the implementation of those technologies in the analysed contexts are done in a simple and unified manner allowing each EU Member State to adopt and use these new technologies in case a disaster occurs within their borders or if a major disaster is likely to hit the European continent. Otherwise, if the unification of an AI and VR/AR system in preventing and managing disasters is not put into practice, most of the available technologies might remain only theoretical thus creating potential damages for the generations to come.

19.10 Conclusions

This chapter has analysed the role of AI and VR/AR in preventing natural hazards. It has also examined the legal framework that surrounds the use of these technologies within the European Union.

It emerged that the EU not only promotes the use of these new technologies (thus recognizing its benefits and technological potential), but also identifies its potential risks: non-transparent decision-making, lack of privacy, discrimination, and gender inequality. In this regard, the European Commission has used a risk-based approach, thus defining a common framework for AI and VR/AR technologies within the European borders.

The chapter has also been focused on framing the complex scenario of the European Disaster Risk Management and Disaster Risk Reduction regulations and policies. For instance, it has been demonstrated that among the European DRM and DRR framework, it is appropriate and recognized the use of artificial intelligence and virtual/augmented reality.

Finally, it has emerged that AI and VR/AR technologies are only partially used by Member States in the prevention of natural hazards because of a lack of uniformity in the discipline regulating their use. For this reason, the future policymaker, in order to make the adoption of these technologies in preventing natural hazards much more effective, will have the task to regulate their use in an equally simple and cost-effective manner.

References

Ahmad Mahmood BA, Qadir AS, Mohammed EA (2022) Tort liability arising from the act of artificial intelligence techniques. Eur Scholar J 3(4):80–87

Auer A (2019) Rechtsfähige Softwareagenten: Ein erfrischender Anachronismus. Verfassungsblog, 30 Sept 2019. https://intr2dok.vifa-recht.de/servlets/MCRFileNodeServlet/mir_derivate_00007429/Rechtsfhige_Softwareagenten_Ein_erfrischender_Anachronismus.pd

Barnier M (2006) For an European civil protection force: Europe aid. https://www.europarl.europa.eu/meetdocs/2004_2009/documents/dv/031006barnier_/031006barnier_en.pdf

Bassyouni Z, Elhajj IH (2021) Augmented reality meets artificial intelligence in robotics: a systematic review. Front Robot AI 8:296. https://doi.org/10.3389/FROBT.2021.724798/BIBTEX

Beck S (2016) The problem of ascribing legal responsibility in the case of robotics. AI Soc 31:473–481

Beckers A and Teubner, G (2021) Three liability regimes for artificial intelligence. Algorithmic Actants, Hybrids, Crowds, Hart, Oxford, pp 3–22

Berchtold C et al (2020) Barriers and facilitators in interorganizational disaster response: identifying examples across Europe. Int J Disaster Risk Sci 11:46–58. https://doi.org/10.1007/s13753-020-00249-y

Boin A (2019) The transboundary crisis: why we are unprepared and the road ahead. J Contingencies Crisis Manag 27:94–99

Braun M, Hummel P, Beck S (2018) Primer on an ethics of AI-based decision support systems in the clinic. J Med Ethics 47(12). https://doi.org/10.1136/medethics-2019-105860

Cauffman C (2018) Robo-liability: the European Union in search of the best way to deal with liability for damage caused by artificial intelligence. Maastricht J Eur Comp Law 25(5):527–532

Cipresso P et al (2018) The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Front Psychol 9:2086. https://doi.org/10.3389/FPSYG.2018.02086/BIBTEX

Dattamajumdar A, Keat W (2021) An early warning AI-powered portable system to reduce workload and inspect environmental damage after natural disasters, 1–16

De La Rosa T, Olaya A. G, Borrajo D (2007). Using cases utility for heuristic planning improvement. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4626 LNAI: 137–148. https://doi.org/10.1007/978-3-540-74141-1_10

DWS, Europe set for record wildfire destruction in 2022, 14 Aug 2022. https://www.dw.com/en/europe-set-for-record-wildfire-destruction-in-2022/a-62802068

Ebers M et al (2021) The European Commission's proposal for an artificial intelligence act—a critical assessment by members of the robotics and AI law society (RAILS). J 4(4):589–603, ì. https://doi.org/10.3390/j4040043

Femia P (2019) Soggetti responsabili. Algoritmi e diritto civile. In: Femia P (ed) Soggetti giuridici digitali. Sullo stato privatistico degli agenti software autonomi, Edizioni Scientifiche Italiane, Napoli, pp 9–10

Galgano F (2007) Tutto il rovescio del diritto. Giuffrè, Milano

Gladden ME (2019) Who will be the members of Society 5.0? Towards an anthropology of technologically posthumanized future societies. Soc Sci 8(5):148. https://doi.org/10.3390/socsci8050148

Herrmann-Pillath C (2021) A Copernican moment: engaging economic ethics in orchestrating the geocentric turn in economics. In: Bohle M, Marone E (eds) Geo-societal narratives. Palgrave Macmillan, Cham, pp 105–126. https://doi.org/10.1007/978-3-030-79028-8_8

Il Resto del Carlino, Alluvione nelle Marche: tutti i dati, ora per ora: ecco cosa è successo, 21 Sept 2022. https://www.ilrestodelcarlino.it/marche/alluvione-cosa-successo-1.8098852

Kankanhalli A et al (2019) IoT and AI for smart government: a research agenda. Gov Inf Q 36:304–309

Kelsen H (1941) The pure theory of law and analytical jurisprudence. Harv Law Rev 55(1):40–70
Li X et al (2018) A critical review of virtual and augmented reality (VR/AR) applications in construction safety. Autom Constr 86:150–162. https://doi.org/10.1016/J.AUTCON.2017.11.003
Lin J, Cao L, Li N (2019) Assessing the influence of repeated exposures and mental stress on human wayfinding performance in indoor environments using virtual reality technology. Adv Eng Inform 39:53–61. https://doi.org/10.1016/J.AEI.2018.11.007
Lin J et al (2020) How occupants respond to building emergencies: a systematic review of behavioral characteristics and behavioral theories. Saf Sci 122:104540. https://doi.org/10.1016/J.SSCI.2019.104540
Pisanò A (2012) Diritti deumanizzati. Animali, ambiente, generazioni future, specie umana, Giuffrè, Milano
Regulation EU 2021/241 of the European Parliament and of the Council of 12 February 2021 establishing the Recovery and Resilience Facility. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R0241&from=IT
Reiners D et al (2021) The combination of artificial intelligence and extended reality: a systematic review. Front Virtual Real 2. https://doi.org/10.3389/FRVIR.2021.721933
Sahu CK, Young C, Rai R (2020) Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review. Int J Prod Res 59(16):4903–4959.https://doi.org/10.1080/00207543.2020.1859636
Scherer MU (2016) Regulating artificial intelligence systems: risks, challenges, competencies, and strategies. Harv J Law Technol 29(2):354–400
Sun R et al (2020) A review of risk analysis methods for natural disasters. Nat Hazards 100(2):571–593. https://doi.org/10.1007/S11069-019-03826-7
Teubner G (2006) Rights of non-humans? Electronic agents and animals as new actors in politics and law. J Law Soc 33(4):497–521
Teubner G (2018) Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten digital personhood? The status of autonomous software agents in private law. Ancilla Iuris 106:108–149
Topildiev B, Khursanov R (2021) Legal fact based on occurrence civil rights and obligations. Ilkogretim Online 20(3):1695–1698. https://doi.org/10.17051/ilkonline.2021.03.195
Van Dijk N (2020) In the hall of masks: contrasting modes of personification. In: Hildebrandt M, O'Hara K (eds) Life and the law in the era of data-driven agency. Edward Elgar, Cheltenham, p 231
Villani S (2017) The EU civil protection mechanism: instrument of response in the event of a disaster. Rev Universitaria Eur 26:121–148
Von Savigny FC (1840) System des heutigen römischen Rechts. Veit und Comp, Berlin
Zhou Y et al (2014) Local spatial and temporal factors influencing population and societal vulnerability to natural disasters. Risk Anal 34(4):614–639. https://doi.org/10.1111/risa.12193

Legal sources

European Commission (a). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM(2021) 206 final, 2021/0106 (COD)), 21 April 2021. https://op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585-01aa75ed71a1/language-en/format-PDF/source-205836026
European Commission Expert Group on Liability and New Technologies—New Technologies Formation. Liability for Artificial Intelligence and Other Emerging Digital Technologies, 27 Nov 2019

European Parliament and Council of the European Union. Directive 2006/42/Ec of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast) (OJ L 157/24). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:157:0024:0086:EN:PDF

European Commission. Proposal for a Regulation of the European Parliament and of the Council on machinery products (COM(2021) 202 final 2021/0105 (COD)), 21 April 2021. https://ec.europa.eu/docsroom/documents/45508

Presidency of the Council of the European Union. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts—Presidency compromise text (2021/0106(COD)), 31 March 2022. https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf

European Commission (b). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 2030 Digital Compass: the European way for the Digital Decade (COM(2021) 118 final), 9 March 2021. https://ec.europa.eu/info/sites/default/files/communication-digital-compass-2030_en.pdf

European Commission. Press release. NextGenerationEU: first annual report on the Recovery and Resilience. Facility finds implementation is well underway, 1 March 2022. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1198

European Commission (c). Communication from the Commission to the European Parliament and the Council on the Third Progress Report on the implementation of the EU Security Union Strategy (COM(2021) 799 final), 8 Dec 2021. https://op.europa.eu/en/publication-detail/-/publication/d5bc5c70-581d-11ec-91ac-01aa75ed71a1/language-en/format-PDF

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Europe's Media in the Digital Decade: An Action Plan to Support Recovery and Transformation COM(2020) 784 final, 3 December 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0784&from=EN

Commission Staff Working Document (2017) Action Plan on the Sendai Framework for Disaster Risk Reduction 2015–2030—A disaster risk-informed approach for all EU policies; SWD(2016) 205 final

Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial Intelligence for Europe (COM(2018) 237 final), 25 April 2018. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN

Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Coordinated Plan on Artificial Intelligence (COM(2018) 795 final), 7 Dec 2018. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:795:FIN

European Environment Agency. Economic losses and fatalities from weather- and climate-related events in Europe, 03 Feb 2022. https://www.eea.europa.eu/publications/economic-losses-and-fatalities-from

European Commission Expert Group on Liability and New Technologies—New Technologies Formation. Liability for Artificial Intelligence and Other Emerging Digital Technologies, 27 Nov 2019. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf

European Commission, Communication from the European Commission to the European Parliament and the Council (COM(2010) 600 final), 26 Oct 2010. https://eur-lex.europa.eu/legal-content/HR/TXT/?uri=CELEX:52010DC0600

European Commission, Decision No 1313/2013/EU of the European Parliament and of the Council of 17 December 2013 on a Union Civil Protection Mechanism Text with EEA relevance. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013D1313

European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 2030 Digital Compass: the European way for the Digital Decade (COM(2021) 118 final), 9 March 2021. https://ec.europa.eu/info/sites/default/files/communication-digital-compass-2030_en.pdf

European Commission. Communication from the Commission to the European Parliament and the Council on the Third Progress Report on the implementation of the EU Security Union Strategy (COM(2021) 799 final), 8 Dec 2021. https://op.europa.eu/en/publication-detail/-/publication/d5bc5c70-581d-11ec-91ac-01aa75ed71a1/language-en/format-PDF

European Commission. Coordinated Plan on Artificial Intelligence 2021 Review (COM(2021) 205 final Annex), 21 April 2021. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

European Commission. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM(2021) 206 final, 2021/0106 (COD)), 21 April 2021. https://op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585-01aa75ed71a1/language-en/format-PDF/source-205836026

European Commission. Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC (COM(2020) 825 final, 2020/0361 (COD)), 15 Dec 2020. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-parliament-and-council-single-market-digital-services-digital-services

European Commission. Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) (COM(2020) 842 final 2020/0374 (COD)), 15 Dec 2020. https://op.europa.eu/en/publication-detail/-/publication/aa57c5e1-46fa-11ec-89db-01aa75ed71a1/language-en

European Commission. Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) (COM(2020) 767 final 2020/0340 (COD)), 25 Nov 2020. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-parliament-and-council-single-market-digital-services-digital-services

European Commission. White Paper on Artificial Intelligence—A European approach to excellence and trust (COM(2020) 65 final), 19 Feb 2020. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

European Forum for Disaster Risk Reduction (2021) Roadmap 2021–2030 For a disaster-resilient European and Central Asian region by 2030. https://efdrr.undrr.org/sites/default/files/2022-01/EFDRR%20Roadmap%202021-2030.pdf

European Parliament. Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 27 Jan 2017. https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html

European Parliament. Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 16 Feb 2017. https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html#:~:text=European%20Parliament%20resolution%20of%2016%20February%202017%20with,on%20Civil%20Law%20Rules%20on%20Robotics%20%282015%2F2103%20%28INL%29%29

European Union, Commission Implementing Decision of 16 October 2014 No 2014/762/EU laying down rules for the implementation of Decision No 1313/2013/EU of the European Parliament and of the Council on a Union Civil Protection Mechanism and repealing Commission Decisions 2004/277/EC, Euratom and 2007/606/EC, Euratom (notified under document C(2014) 7489). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32014D0762

European Union, Consolidated version of the Treaty on the Functioning of the European Union, 13 December 2007, 2008/C 115/01. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT

European Union, Next Generation EU. https://next-generation-eu.europa.eu/index_en

High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy Artificial Intelligence (2019) 8 April 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

High-Level Expert Group on Artificial Intelligence. Policy and Investment Recommendations for Trustworthy AI, 26 June 2019b. https://www.europarl.europa.eu/italy/resource/static/files/import/intelligenza_artificiale_30_aprile/ai-hleg_policy-and-investment-recommendations.pdf

UN-EU, Improving Lives Partnership between the UN and EU (2010). https://www.unbrussels.org/report2010/#/page/-1

UNISDR (United Nations International Strategy for Disaster Reduction) (2015) Sendai framework for disaster risk reduction 2015–2030. http://www.wcdrr.org/uploads/Sendai_Framework_for_Disaster_Risk_Reduction_2015-2030.pdf

United Nations General Assembly 2015 Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N15/291/89/PDF/N1529189.pdf?OpenElement

United Nations Office for Disaster Risk Reduction, (2020) Europe's opportunity to manage risk and build resilience Recommendations to the European Green Deal, 1–11

Concluding Remarks

Vladimir Geroimenko

This research monograph provides a broad view of the areas, approaches and applications related to combining the potentials of Augmented Reality (AR) and Artificial Intelligence (AI) technologies. These two technologies have been rapidly advancing in recent years, and they are able to offer new creative solutions and experiences by supporting each other. AR is capable of supporting AI and AI can also support AR, and consequently these technologies can be fused together and serve as co-technologies, establishing a new emerging field of research and applications that incorporate both AR and AI.

The book considers several practical, theoretical and cultural aspects of integrating Augmented Reality and Artificial Intelligence, the educational use of intelligent augmented environments, the creative and productive fusion of the two powerful technologies in medicine, healthcare and physical activity, and combining AR and AI to enhance services, retail and recommendations. The monograph addresses an extensive range of particular and special topics, such as augmenting visual information using machine learning models, AI-based data analysis for AR user's experience, immersive intelligent tutoring systems, AI-powered "augmented" dentistry, AI-based services for generating AR experiences, smart AR in metaverse-tailing and luxury fashion retail, the benefits of the AR-AI technological alliance in dental education, historical reconstructions, pharmacy, literary fiction, preventing natural disasters, AI-simulated life-like clinical examinations, physical activities and exergames. The levels and methods of the consideration are various, from practical applications and systematic analyses to phenomenological and speculative approaches.

In other words, the book shows how AR and AI can work together producing both AI-empowered AR applications and AR-enabled AI applications in the framework of the coherent integration of AI and AR. This powerful fusion of the two cutting-edge technologies has an enormous potential for current and future developments in

V. Geroimenko (ed.), *Augmented Reality and Artificial Intelligence*, Springer Series on Cultural Computing,
https://doi.org/10.1007/978-3-031-27166-3

a vast variety of areas. As these key technologies continue to evolve and improve, their combination will constantly enable new possibilities; for example, it might help to shape the metaverse and to improve the way people interact with computers in the ground-breaking metaverse environment in the years to come.