María José Abásolo
Ailyn Febles Estrada
Carlos De Castro Lozano (Eds.)

Communications in Computer and Information Science 2140

Applications and Usability of Interactive TV

12th Iberoamerican Conference, jAUTI 2023
Havana, Cuba, October 16–17, 2023
Revised Selected Papers

Springer

Communications in Computer and Information Science 2140

Rationale
The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics
The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors
Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (http://link.springer.com) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at http://link.springer.com/bookseries/7899. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process
The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing
CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start
To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

María José Abásolo · Ailyn Febles Estrada ·
Carlos De Castro Lozano
Editors

Applications and Usability of Interactive TV

12th Iberoamerican Conference, jAUTI 2023
Havana, Cuba, October 16–17, 2023
Revised Selected Papers

Springer

Editors
María José Abásolo
National University of La Plata
Buenos Aires, Argentina

Ailyn Febles Estrada
Universidad de Las Ciencias Informáticas
Havana, Cuba

Carlos De Castro Lozano
University of Córdoba
Córdoba, Spain

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-70438-3 ISBN 978-3-031-70439-0 (eBook)
https://doi.org/10.1007/978-3-031-70439-0

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

Preface

The XII Ibero-American Conference on Applications and Usability of Interactive Digital Television (jAUTI 2023) was jointly organized by the Union of Cuban IT Professionals (Cuba) and RedAUTI (Thematic Network on Applications and Usability of Interactive Digital Television). This conference took place from October 16–17, 2023, in the city of La Habana (Cuba). This book contains a collection of 9 papers referring to the design, development, and user experiences of applications for Interactive Digital Television and related technologies, that were selected as the best papers from 25 papers received at the event, after a two-round peer-review process.

June 2024

María José Abásolo
Carlos De Castro Lozano

Organization

Program Committee Chairs

Ailyn Febles Estrada	University of Computer Sciences, Cuba
Carlos De Castro Lozano	University of Córdoba, Spain
María José Abásolo	National University of La Plata, Argentina

Program Committee Members

Alan Guedes	University of Reading, UK
Alcina Prata	Polytechnic Institute of Setúbal, Portugal
Ana Velhinho	Digimedia University of Aveiro, Portugal
Anelise Jantsch	Federal University of Rio Grande do Sul, Brazil
Antoni Oliver	University of the Balearic Islands, Spain
Beatriz Sainz de Abajo	University of Valladolid, Spain
Cecilia Sanz	National University of La Plata, Argentina
Diego Villamarín	Universidad de las Fuerzas Armadas ESPE, Ecuador
Fernando Boronat	Polytechnic University of Valencia, Spain
Fernando Fuente Alba	Catholic University of the Holy Conception, Chile
Francisco Montero Simarro	University of Castilla-La Mancha, Spain
Gonzalo Olmedo Cifuentes	Universidad de las Fuerzas Armadas ESPE, Ecuador
Joaquín Danilo Pina Amargós	CUJAE, Cuba
Jorge Abreu	Digimedia University of Aveiro, Portugal
José Luis Arciniegas Herrera	University of Cauca, Colombia
José Maria Buades Rubio	University of the Balearic Islands, Spain
Manuel González Hidalgo	University of the Balearic Islands, Spain
Patrícia Oliveira	Digimedia University of Aveiro, Portugal
Pedro Almeida	Digimedia University of Aveiro, Portugal
Raisa Socorro Llanes	CUJAE, Cuba
Rita Oliveira	Digimedia University of Aveiro, Portugal
Rita Santos	Digimedia University of Aveiro, Portugal
Sandra Baldassarri	University of Zaragoza, Spain
Telmo Silva	Digimedia University of Aveiro, Portugal
Vagner Beserra	University of Tarapacá, Chile
Valdecir Becker	Federal University of Paraíba, Brazil

Contents

Technologies, Services and Content Creation for Interactive Digital TV

Architecture for Interactivity in DTV Adapted to Cuban Conditions

Ariel Alfonso Fernández Santana, Joaquín Danilo Pina Amargós(✉), and Raisa Socorro Llanes

Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE), Havana, Cuba
jpina@ceis.cujae.edu.cu

Abstract. Digital Television (DTV) allows the incorporation and transmission of data that is interpreted by decoder boxes and that the end user perceives with a certain degree of interactivity. Currently, most of the services and technologies used to achieve interactivity in DTV (DTVi) are limited by the charging of high rates and their source code is not available to adapt their operation to local conditions. Among the interactive applications are video games and virtual visits, which contribute to making the use of television more attractive and educational. The present research proposes an architecture adapted to the technological conditions of Cuba, which contributes to incorporating advanced interactivity functionalities in DTV. To do this, an analysis of the available technologies is carried out, opting for a free software solution that allows the creation of these applications and their integration into the same system following an architecture based on microservices and the HbbTV (Hybrid Broadcast Broadband Television) standard. The prototypes were evaluated on various low-performance hardware setups, simulating the proposed DTV hardware for Cuba, demonstrating satisfactory performance in terms of latency, software, and hardware efficiency. With the development of research, we contribute to enhancing the country's technological heritage. Since television is a widely used broadcast medium in today's society, its implementation is reflected as a clear benefit to different sectors. In addition, it promotes the growth of the country's technological sovereignty by being a reference in the development of video games, virtual visits, or other types of interactive content for DTV, since it creates a characterization of the main attributes that these applications must have.

Keywords: Interactive digital television · Software architecture · Free software

1 Introduction

DTV (Digital Television) is known as a set of audiovisual transmission and reception technologies that use digital signals instead of the traditional analogue signals of television.

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 3–17, 2024.
https://doi.org/10.1007/978-3-031-70439-0_1

As stated in [2] and [19], DTV presents interesting innovations compared to traditional TV, such as: various transmission formats, simultaneous transmissions and interactivity.

Based on [15] it can be stated that interactivity allows television channels to offer a set of additional services by incorporating advanced communication, participation, and social services functions for the development of computerization. This interactivity is classified into two levels: local interactivity and return channel interactivity.

Interactive Digital Television (DTVi) is a growing emerging scenario that offers the possibility of transmitting informative and recreational content using this medium with great penetration in most towns in the world. Among the most striking contents is interactive multimedia that includes video games and virtual tours. Together they offer a truly integrated experience combining television, computers, the publishing industry and telecommunications.

An interactive multimedia product allows the user to initiate and develop a dialogue, consult information, explore, discover and acquire new knowledge; Therefore, they highlight the great benefits that its use provides in the field of education and entertainment. It is considered a technology that promotes creativity through computing systems. The production and creation of virtual systems reduces the waste of technical and economic resources. It focuses on the use of available technological resources, their advances, and multimedia tools to develop interactive, simple products, in which, using various design and creativity techniques, a lot of informative content can be included. Furthermore, an interaction that requires the user to facilitate attention, understanding, and retention of information in an intuitive and spontaneous way.

Among the various products in multimedia format that exist, three important collections of educational software stand out in Cuba: the "Multisaber", "Navegante" and "Futuro" collections [17] collection, aimed at transmitting knowledge in a more entertaining way to students. primary, secondary and higher education students respectively. Systems like these constitute a valuable teaching-learning medium, which offers varied perspectives of projection to the teaching-educational process. Some provide a curricular and multidisciplinary approach due to their relationship with the contents of the programs of each subject, and others contribute to the formation of a comprehensive general culture. However, they are not prepared to be deployed in the mass environment of DTV, a way in which it could be delivered to homes regardless of their purchasing level.

Currently, DTVi is mainly used to transmit multimedia content that is controlled by companies and corporations that respond to their own interests above those of the users. Regarding Cuban Television, there are certain technological and economic conditions that have a notable influence when it comes to implementing interactivity. A first point is that only local interactivity is possible due to the absence of a return channel, which makes bidirectional communication inapplicable [18]. Furthermore, the decoder boxes available at the country level for use by the population are of low performance.

On the other hand, useful content regarding the topic and relevant technologies is dispersed on the Internet, making it difficult for those interested in

developing applications in this environment to easily assimilate it. Acquiring a similar technology from a foreign company would imply a significant economic expense, since it would not be viable due to the blockade imposed by the United States government against Cuba [9]. At the same time, the proprietary technologies with which they are developed do not allow existing content to be adapted to the low-performance environments that the country has, nor to integrate them with other solutions.

This article describes the state of the art regarding interactivity in DTV and existing solutions that contribute to the coexistence between different television standards. Next, to achieve the desired interactivity, it is proposed to use a system appropriate to the HbbTV (*Hybrid Broadcast Broadband TV*) standard that works in conjunction with current broadcasting solutions and the results obtained in the 360 virtual visit scenarios are presented. Educational video games where it was validated. Access to said system will be through a second screen (external device) such as a smartphone, tablet or using a smart television or digital TV decoder that has Internet availability, allowing client-server communication.

Taking into account the above, the research is based on proposing an architecture adapted to the technological conditions of Cuba, which contributes to incorporating advanced interactivity functionalities in Cuban DTV.

The development of the topic that is raised would contribute to enhancing the technological heritage of the country. Since television is a widely used broadcast medium in today's society, its future implementation will be reflected as a clear benefit to different sectors of society. In addition, it will provide technological sovereignty and may become a reference in the development of multimedia, video games, virtual visits, or other types of interactive content for DTV, since a characterization of the main attributes that these applications must have will be created.

2 Previous Works

IBB (Integrated broadcast-broadband) systems have proven to be a valid solution to allow broadcasters to bring new types of emerging content to the end user (ITU-R). These systems combine traditional services with multifunctional applications, similar to Web-based services, which is why many European broadcasters have begun to take advantage of these new opportunities. For example, there is the case of the BBC (British Broadcasting Corporation) which launched a new testing system that helps manufacturers verify the functionalities and performance of HbbTV devices so that television applications and second screen services sync accurately with a program or channel being viewed on the TV [11].

On the other hand, the HbbTV Certification Group developed an HbbTV Operator Application (OpApp) to allow television broadcasters to control the user experience on devices such as Set-Top Box (STB) and Smart TV operating across different ecosystems and devices, suitable for any distribution medium (e.g. cable, terrestrial, and satellite) [13].

It is important to note that IBB is usually associated with only one DTV standard, resulting in unique systems [8]. To overcome this limitation, several studies have proposed solutions that allow coexistence between different TV standards. In [19] a hybrid TV system is presented that combines ISDB-T and HbbTV, supporting different types of synchronization mechanisms between the ISDB-T signal and the associated HbbTV applications. In [6] the authors present a hybrid network for the provision of linear services on portable and mobile devices, based on the cooperation between the *broadcast* system and the *broadband*. [3] describes a platform to distribute end-to-end hybrid multimedia content (*end-to-end*) in a multi-device scenario compatible with HbbTV, where it allows the use of live streaming via HTTP (*HyperText Transfer Protocol*) and real-time transport protocols (*Real-time Transport Protocols*) that are not supported by the HbbTV standard. Another study [10] takes advantage of terrestrial and satellite backhaul in combination with Internet connectivity to expand the reach of television broadcasts to devices such as *smartphones* and *tablets*. These investigations demonstrate that it is possible to create a system to add interactivity functions that works in conjunction with existing broadcasting solutions, highlighting usability, accessibility, intuitive interaction and immediate feedback as the main characteristics of interactive applications.

In [18] a new software solution is presented that demonstrates the possibilities of DTVi in a real scenario. The solution called TVC+ collects useful information available on the Internet and integrates it with DTVi services. Some of its functionalities have already been deployed in various scenarios [1], demonstrating its usefulness in some areas of the 2030 Sustainable Development goals of the UN (United Nations Organization): Education, Health, Food, and Heritage [4]. TVC+ allows the transmission of content in the DTVi standard that is required.

There are several technical standards that regulate the operation of an interactive television system. Based on the research carried out in [5] and [7], the HbbTV standard stands out for its multiple advantages over other technologies, making it ideal for the development of interactive applications in DTV.

HbbTV is based on a set of existing open standards, which define how to interact with multimedia content: OIPF-DAE (*Open Internet Protocol Television - Declarative Application Environment*), CEA (*Consumer Electronics Association*), DVB (*Digital Video Broadcasting*) and W3C (*World Wide Web Consortium*). It goes beyond traditional standards, as it defines its own for graphical interfaces. The OIPF-DAE standard defines JavaScript APIs (*Application Programming Interface*) for television environments, as well as establishes modifications to the CE-HTML language (*Consumer Electronics Association - HyperText Markup Language*) for the creation of interfaces graphics. CEA defines the APIs for on-demand services, as well as access to UpnP (*Universal Plug and Play*) networks and the Internet. The DVB standard defines the transport and signaling layer of interactive content and W3C defines the Web standards: HTML (*HyperText Markup Language*), CSS (*Cascading Style Sheets* 3), JavaScript, DOM (*Document Object Model*), among others for the presentation of interactive content. With these standards, HbbTV makes content from different providers and even different transmission media accessible through the

same interface and can be processed on low-performance devices. In turn, it uses the MPEG-DASH (*Dynamic Adaptive Streaming over HTTP*) specification in order to provide support for adaptive streaming according to the client's capabilities through the HTTP protocol. Adaptive bitrate content transmission is one of the essential features of HbbTV achieving a better user experience on high saturation or low speed networks [14].

Several years after the standard was published, numerous organizations have created solutions to generally integrate them with their hardware products. Such is the case of Samsung, LG and Sony that have combined their Smart TV systems with HbbTV. Various television providers, especially in Europe, have begun to transmit content under the HbbTV standard. Currently, there are more than 30 countries where the standard has been deployed, highlighting, in addition to the European market, Russia, Australia, and New Zealand [12]. On the other hand, government organizations, together with these suppliers, have created recommendations and specifications with the characteristics that receiving equipment and software products must have to be sold in their countries.

3 Proposed Architecture

The proposed architecture to allow the incorporation of interactivity in Cuban DTV is based on the HbbTV standard; this decision is based on all the advantages mentioned above.

Figure 1 shows a diagram of how the proposed architecture is planned to be deployed.

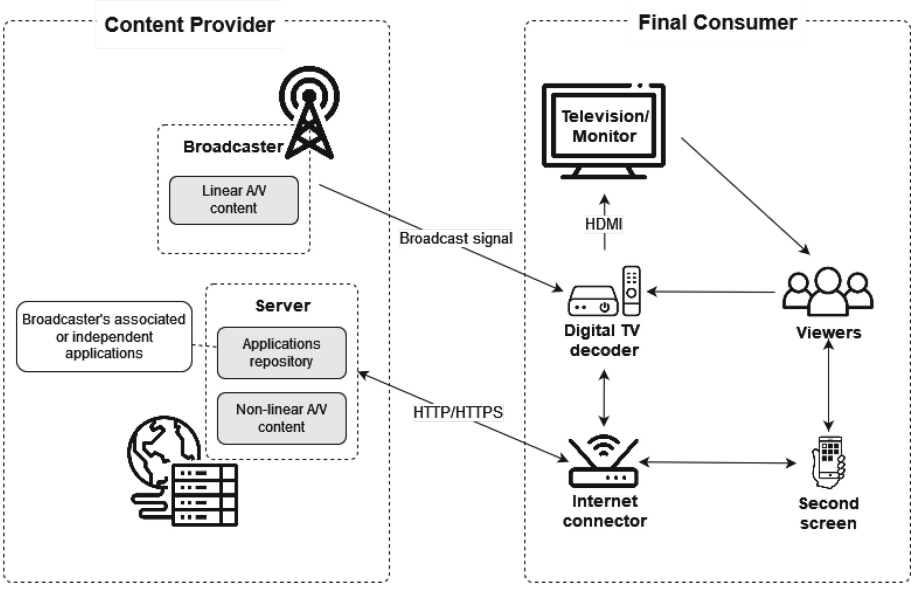


Fig. 1. Proposal for the deployment of the architecture.

In the area corresponding to the **content provider**, the parts referring to the **Broadcaster** are located, where the **linear A/V (audio or video) content** is stored, and the **applications' repository** and **non-linear A/V content**. Therefore, it can be deduced that the existence of one or more servers will be necessary, responsible for storing both the multimedia content and the applications available through the Internet.

Additionally, this part of the platform includes the necessary elements so that the **final consumer** can access all the necessary information in order to be able to acquire and reproduce complementary content or acquire the information associated with the programs that are available they transmit. Therefore, you must have a **digital TV decoder**, an **internet connector** (such as a router or similar equipment with internet access) and a **television or monitor**. Optionally, you can have some type of **second screen** (such as *smartphone*, *tablet*, computer or similar).

As can be seen, the **Broadcaster** is responsible for transmitting the **linear A/V content** through the *broadcast* signal, which is received by the **digital TV decoder** to stream the content to the **TV or monitor**. **Viewers** interact with the **digital TV decoder** to select the **linear A/V content** they wish to consume and view it through the **TV or monitor**.

The **end consumer** has two ways to access **non-linear A/V content** and **applications**. One is by having a **digital TV decoder** that has a network interface that allows it to connect to the **internet connector**, whether wireless or wired. And another, using a **second screen** that connects in the same way to the **internet connector**.

The applications hosted in the **applications' repository** can be associated with the **Broadcaster** (applications that are part of the DTV IBB services) or independent (applications that are not available through the DTV service).

The return channel is established through the internet, as proposed by HbbTV. So, **viewers** can consume **linear A/V content** and interact with **non-linear A/V content** and **advanced interactivity applications** through the connection established between the **internet connector** and the **Server**.

It should be noted that these devices may or may not be integrated, for example, the **second screen** may have an Internet connection, as may the **digital TV decoder**, just as the TV may have the decoder integrated with or without an Internet connection. Depending on their particular conditions, each user will personalize the way they consume the interactive services.

The applications and their content will be downloaded on the **second screen** or on the **digital TV decoder** with Internet access and can be executed simultaneously with the transmission of the program or after its broadcast. Take a cooking show as an example: you can check the recipe on the **second screen** during or after the show is broadcast. The same would apply to a virtual visit to a place of interest, with its projection in panoramic view or video game format during or after its broadcast. Additionally, in certain programs, the interactivity application can be used to vote, make comments or questions, and even, during a

virtual class, carry out an evaluation of the class content. This allows the viewer to submit their response and receive immediate feedback.

Taking into account that the system is required to have a high level of modularity and scalability, since it is essential that it allows the increase of its functionalities, it was decided to use an architecture based on microservices. This brings with it a series of advantages explained below. In principle, it contributes to decomposing the application into different services, with the aim of obtaining high availability, low coupling, decentralization and fault tolerance. Another advantage that stands out with respect to the scalability of the system is that because each microservice works independently of the rest, the process of adding, removing or modifying the modules that make up the application is simplified. This allows a first approach to the implementation of this solution to be achieved as the components that make up the system are developed, without the need to put it into operation for all the desired scenarios from the beginning. On the other hand, they allow the solution to not be anchored to a particular technology, which encourages its evolution as new technologies emerge without affecting the deployment of the system.

At the same time, to address the problems that arise with the use of this type of architecture, taking into account the management of communication between microservices and their monitoring, it is proposed to follow the design patterns *API Gateway* and *Saga*.

The *API Gateway* pattern creates a single entry point for all clients or external applications that handles requests by redirecting them to the appropriate service or services. This practice facilitates its integration with other systems and allows services to be changed and reorganized without affecting client applications, since it isolates them from knowing the structure of the application.

For its part, the *Saga* pattern is aimed at solving the challenge of microservices sharing data and communicating with each other, because each microservice runs in isolation and with its own database. Its implementation in the system proposes having a transaction manager to manage and guarantee the integrity and consistency of data between microservices. Each transaction publishes an event or message when it updates the database, an action that triggers the next step of the transaction and, if any step fails, the ability to undo the preceding transactions.

As mentioned, HbbTV defines Web standards for the presentation of interactive content through W3C. W3C's mission is to bring the *World Wide Web* to its full potential by developing protocols and guidelines that ensure its long-term growth, developing open specifications to improve the interoperability of Web-related products [20]. This implies that the technologies used to develop the interactivity modules of the system must respond to this standard, which is achieved with the use of DOM3 and markup languages such as HTML5, XML (*Extensible Markup Language*) or XHTML (*EXtensible HyperText Markup Language*) assisted with CSS3 and *JavaScript*; while for operations such as rendering and graphic calculations on the GPU (*Graphics Processing Unit*) it validates the use of WebGPU (*Web Graphics Processing Unit*) and WebGL (*Web Graphics*

Library). Likewise, attention must be paid to ensuring that they do not demand a high capacity of computing resources. Based on these criteria, a technological selection can be made that adapts to the proposed architecture.

With the objective of obtaining early feedback regarding the validation of the proposed architecture, this article focuses on two possible scenarios to which it can be applied to achieve interactivity in Cuban DTV.

In a study conducted at [21] during the spread of COVID-19, on how 360 virtual visits can reduce people's psychological stress, the results indicated a decrease in stress and an increase in enjoyment when participating in this way of virtual reality. Similarly, in [16] it was proven that playing video games can positively influence reducing stress and anxiety in children, adults and older adults, in addition to being a potential source of knowledge depending on the particular characteristics of the game in question. Consequently, it was decided that the proposed architecture will be validated in the scenario of 360 virtual tours and video games.

Therefore, an analysis was carried out of the fundamental frameworks for the development of 360 virtual visits in an environment conducive to DTV, of which Three.js stood out for graphical management on the client side. The choice is based on the fact that it is a lightweight JavaScript library that has mechanisms for resource optimization, an important advantage in the work environment. In addition, it is constantly updated with new features, stability improvements and bug fixes; which guarantees technical support. For the development of the rest of the client-side application, React.js was used, since using a virtual DOM allows great performance and largely uses JavaScript functionalities.

Similarly, an analysis of the technologies available for video game development was carried out, highlighting the Pixi.js JavaScript library. The decision is based on the fact that it is completely free, it needs few requirements for its implementation and, as it is multi-platform, it is compatible with the different operating systems of Smart TVs and STBs.

In the case of the server-side application, it was decided to use Node.js with Nest.js based on the fact that it consumes considerably fewer system resources compared to other options (Asp.Net or Spring), provides a default architecture, uses JavaScript, and It makes testing easier as it allows you to develop applications using a small amount of code.

Figures 2 and 3 represent the proposals to implement the API Gateway and Saga pattern respectively in the 360 virtual tour application. In the case of the pattern, Saga uses RabbitMQ as an intermediary between the events that the services publish when updating the database.

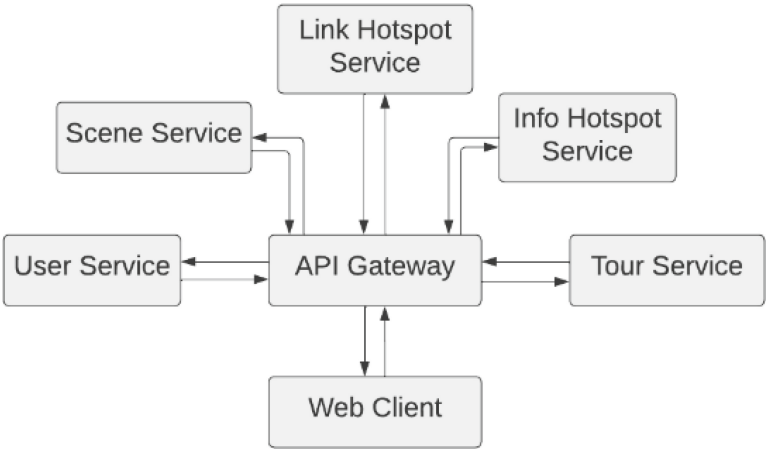


Fig. 2. *API Gateway* design pattern for the 360 virtual tours application.

Figure 4 shows the system represented in a layered structuring architecture with a reuse approach where you can see the different packages that make up the solution. In the specific and general layer, the particular packages of the application are shown, the views with which the user will interact, the models that include the images and resources and the general logic of the business along with the tools used; These are the least reusable components. In the intermediate layer are the libraries and frameworks used as a complement in the development of the system, created by third parties, and which can be reused for the development of similar applications. Finally, the software layer of the system contains the components of the operating system, such as the protocols that manage access to data and resources.

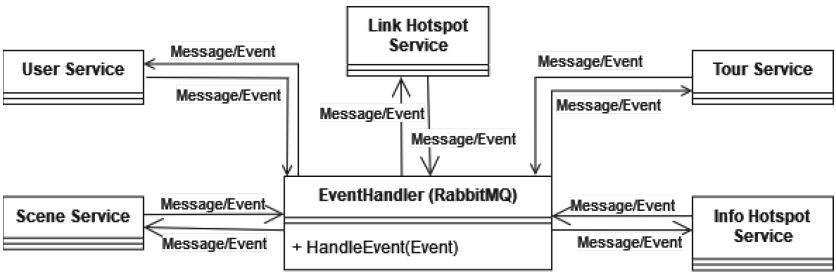


Fig. 3. *Saga* design pattern for the 360 virtual tours application.

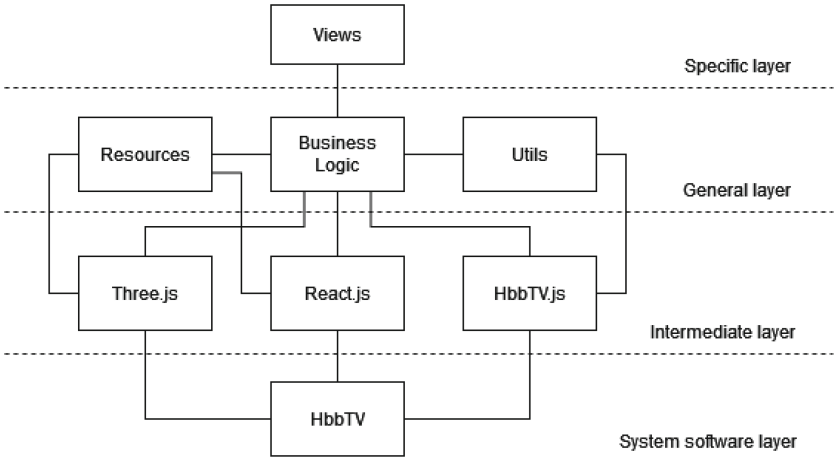


Fig. 4. Layered structuring architecture with a reuse approach.

4 Results Analysis

To validate the proposed architecture in the indicated scenarios, functional prototypes were developed that would allow its deployment to be verified in an environment close to the real one using the selected technology. Once the prototypes are developed, they will be tested on desktop, mobile and SBC (*Single-Board Computer*) computers. Mozilla Firefox, Google Chrome and Opera were used as Web browsers.

4.1 360 Virtual Tours

In the scenario of 360 virtual tours, a system was developed that allows the formation and visualization of these. Figure 5 shows screenshots of its operation, both of the editor and the viewer. On the left you can see the start menu, where the virtual tours to which the user has access and that can be viewed, modified or deleted are listed. From this view, you can also create a new virtual tour. The editor and viewer are shown on the right. The user can create the virtual tour from the editor, where the scenes and points of interest that make it up are created. While to visualize a virtual tour, a view is used that encourages the user's immersion in the tour, while allowing interaction with the elements of the environment.

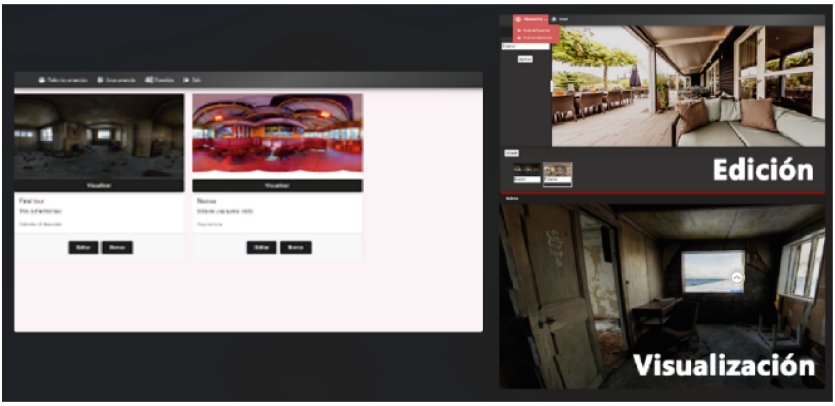


Fig. 5. System for the formation and visualization of 360 virtual tours.

4.2 Videogames

In order to promote learning through video games, it was decided that the prototype to be developed was based on logic and knowledge with a question and answer approach. Based on this, blocks of questions grouped by subject were created that the user must answer correctly to win.

Figure 6 shows the view where the questions are shown, and the user makes their choice. Each game begins with 3 lives represented by hearts. When the player answers a question incorrectly, they will lose a heart, which when reduced to 0 ends the game. If all the questions of the subject can be answered without losing all the hearts, the subject is defeated.

4.3 Evaluation Results

To assess the feasibility of the proposed architecture, extensive evaluations were carried out on the developed prototypes. These evaluations included a series of black box approach tests, to verify they met the stated requirements, which yielded satisfactory results and performance tests on various low-performance hardware setups, simulating the proposed hardware configuration for Cuban DTV.

Fig. 6. Screen for user interactivity of an educational video game.

The prototypes were tested on desktop, mobile, and Single-Board Computer (SBC) setups. Performance metrics such as CPU and memory usage were monitored. The results showed that the prototypes performed efficiently within the hardware constraints, with CPU usage averaging around 50% and memory usage peaking at 350 MB, which is within acceptable limits for low-performance devices commonly used in Cuba.

Quality of service parameters such as latency, jitter, and packet loss were measured during the tests in a local network. The latency remained stable, with minimal jitter and no packet loss observed. These results confirm the robustness of the proposed solution in maintaining a high level of interactivity without compromising on user experience. Extra tests should be performed in the deployment phase to measure how many concurrent users the system could handle and how latency will behave based on the server setup and connection infrastructure without significant degradation in performance.

Specific tests were conducted using hardware with similar setup to the proposed DTV hardware setup in Cuba, focusing on the ability to support advanced interactive functionalities like 360 virtual tours and video games. The results indicated that the DTV hardware could adequately support these functionalities, with performance metrics similar to those observed on computers. This demonstrates the feasibility of deploying the proposed solution on the DTV infrastructure in Cuba.

With the development of these applications following the architecture proposed in this research, it is satisfactorily demonstrated how advanced interactivity can be implemented in Cuban DTV in the scenarios of 360 virtual visits and video games. Although research must continue and address other scenarios such as working with a second screen or more complex video games, this first

approach represents a notable step forward in the development of this type of applications since it lays a foundation that developers can follow.

5 Conclusions

The study of the state of the art demonstrated the need to have an architecture adapted to the technological conditions of the country that contributes to incorporating advanced interactivity functionalities in Cuban DTV. This research proposes this architecture and develops prototypes to verify its correct operation in the scenarios of 360 virtual tours and video games.

The proposed architecture is based on the HbbTV standard, so the technologies to be used to implement interactivity must adapt to the W3C and be low impact at the resource level. After analyzing the available technologies, it was decided to develop the client application in charge of managing interactivity with Three.js for 360 virtual tours and Pixi.js for video games, using Node.js and Nest.js on the server side.

The validated scenarios demonstrate the feasibility of the proposed solution and, by using a microservices-based architecture, low coupling between system components is guaranteed. This makes it easier to add new functionalities while reusing already created elements, thus highlighting the benefits obtained. At the same time, it provides the system with the necessary flexibility to replace the frameworks used to implement interactivity as technological development progresses and new options emerge, without affecting other modules of the application; thus contributing to its scalability and continuous improvement.

The evaluation results confirm the viability of the proposed architecture in the Cuban context, showing that the proposed DTV hardware can support the advanced interactive functionalities tested. Consequently, this study represents a contribution to the development of DTVi in Cuba, since a foundation is laid that serves as a guide and support to develop interactive applications in this environment using free software. To continue the work, it is necessary to validate the proposal in real scenarios when the conditions are created.

Acknowledgements. This research has been supported by the Pérez-Guerrero Trust Fund for South-South Cooperation (PGTF) of the United Nations Development Program (UNDP) project INT/19/K08 and the Ministry of Science, Technology and Environment of Cuba (CITMA) DTVi project NPN223LH006-005.

References

1. Abásolo Guerrero, M.J., et al.: La televisión digital interactiva para el mejoramiento de los pueblos latinoamericanos. In: XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja) (2021)
2. Amador-González, M.: Módulo para la extracción de información de fuentes externas para la conformación de noticias en la televisión digital en Cuba. Ph.D. thesis, Bachelor thesis report Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE) (2018)

3. Boronat, F., Marfil, D., Montagud, M., Pastor, J.: HbbTV-compliant platform for hybrid media delivery and synchronization on single- and multi-device scenarios. IEEE Trans. Broadcast. **64**(3), 721–746 (2018). https://doi.org/10.1109/TBC.2017.2781124
4. Desa, U., et al.: Transforming our world: the 2030 agenda for sustainable development (2016)
5. Eslava Arce, I.: Estudio del estándar de televisión digital interactiva hbbtv e implementación de aplicación final (2014)
6. Fam, P.A., Paquelet, S., Crussière, M., Hélard, J.F., Brétillon, P.: Analytical derivation and optimization of a hybrid unicast-broadcast network for linear services. IEEE Trans. Broadcast. **62**(4), 890–902 (2016)
7. Frómeta-García, R.A.: Diseño y desarrollo de videojuego de corte educativo en entorno interactivo de TV digital. Ph.D. thesis, Bachelor thesis report Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE) (2021)
8. Gavrila, C., Popescu, V., Fadda, M., Anedda, M., Murroni, M.: On the suitability of HbbTV for unified smart home experience. IEEE Trans. Broadcast. **67**(1), 253–262 (2020)
9. Ginsburg, M.: Economic and media war against socialist societies. Int. J. Cuban Stud. **14**(2), 272–308 (2022)
10. Hammershoj, A., Nowak, A., Hansen, J.K., Stefanovic, C.: The next-generation television broadcasting test platform in copenhagen. In: 2020 13th CMI Conference on Cybersecurity and Privacy (CMI)-Digital Transformation-Potentials and Challenges (51275), pp. 1–6. IEEE (2020)
11. Hammond, M.: Release of HBBTV/DVB companion synchronisation tools and streams. BBC R&D (2017). https://www.bbc.co.uk/rd/blog/2017-05-second-screen-streams-tools-companion
12. HbbTV: HbbTV deployment countries (2022). https://www.hbbtv.org/deployments/
13. HbbTV-Certification-Group: The benefit of hbbtvopapp for operators and vertical models - the open standard for a unified TV experience across different platforms (2018). https://www.hbbtv.org/wp-content/uploads/2018/09/HbbTV-MG-00632-003-WhitepaperHbbTVOpApp-v1.1.pdf
14. Jakšić, B., Milošević, I., Petrović, M., Ilić, S., Bojanić, S., Vasić, S.: Characteristics of hybrid broadcast broadband television (HbbTV). Bull. Nat. Sci. Res. **7**(1) (2017)
15. Millo Sánchez, R., Morell Pérez, C., García González, C., Siles Siles, I.: La interactividad en la televisión digital: su desarrollo en cuba. Revista Cubana de Ciencias Informáticas **12**(1), 205–219 (2018)
16. Pallavicini, F., Pepe, A., Mantovani, F.: Commercial off-the-shelf video games for reducing stress and anxiety: systematic review. JMIR Ment. Health **8**(8), e28150 (2021)
17. de la Peña Sarracén, E.: El módulo juegos de la colección el navegante en su versión multiplataforma. In: [2019-MADRID] Congreso Internacional de Tecnología, Ciencia y Sociedad (2019)
18. Pina-Amargós, J., Alvarez-Goenaga, D., Villarroel-Ramos, D., Amador-Gonzalez, M., Socorro-Llanes, R.: New functionalities of digital terrestrial television in Cuba to contribute to the informatization of society. Revista Cubana de Ciencias Informaticas **12**, 158–172 (2018)

19. Sotelo, R., Joskowicz, J., Rondán, N.: An integrated broadcast-broadband system that merges ISDB-T with HbbTV 2.0. IEEE Trans. Broadcast. **64**(3), 709–720 (2018)
20. W3C: World wide web consortium (2021)
21. Yang, T., Lai, I.K.W., Fan, Z.B., Mo, Q.M.: The impact of a 360 virtual tour on the reduction of psychological stress caused by covid-19. Technol. Soc. **64**, 101514 (2021)

Development of a Gateway Server to Enhance Early Warning in Digital Terrestrial Television

Gonzalo Olmedo(✉) and Alejandro Salas

WiCOM-Energy Research Group, Universidad de Las Fuerzas Armadas ESPE, Sangolquí, Ecuador
{gfolmedo,jlsalas}@espe.edu.ec

Abstract. This research addresses the analysis, design, development, and implementation of a Gateway Server aimed at decoding and retransmitting early warning messages within the context of digital terrestrial television services. It is crucial to highlight that Ecuador is a disaster-prone area susceptible to natural calamities. The Emergency Warning Broadcasting System is highly beneficial in disseminating information. However, due to its low popularity and certain limitations, developing a Gateway server becomes essential to notify users or clients through various mobile and desktop applications. The system is designed to be versatile and compatible with multiple platforms, including desktop and mobile devices, utilizing iOS and Android native services for notifications. Furthermore, a public server hosted on Digital Ocean, a user-friendly platform offering accessible Cloud Computing services, has been implemented. This server manages the logic related to users, tokens, device identifiers, and alert messages. These services result in a Gateway Server capable of interpreting early warning messages and administering alert distribution through APNs and Google FCM for mobile clients and through direct API consumption toward the server for desktop clients.

Keywords: EWBS · ISDB-T · Gateway Server · TDT

1 Introduction

Disaster prevention and the swift dissemination of early warnings are crucial aspects for the safety and well-being of the population. Throughout history, we have witnessed numerous devastating events, such as the earthquake that struck Tokyo in September 1983, resulting in the loss of 100,000 lives [1]. This tragic event gave rise to Disaster Prevention Day in Japan, commemorated on the first of September. To stress the importance of early warning and disaster prevention, Japan implemented the Emergency Warning Broadcasting System (EWBS) in 1985, marking the beginning of a new era in emergency response [2].

In the international context, natural disasters, such as the devastating earthquake that occurred in Ecuador in April 2016, reaching a magnitude of 7.8 on the Richter scale and resulting in the tragic loss of 655 lives, remind us of the importance of early warning and the need to expand the dissemination of alerts to a broader audience. In this regard,

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 18–30, 2024.
https://doi.org/10.1007/978-3-031-70439-0_2

technological solutions have been sought, such as the implementation of alert systems in the IP telephony network of the Universidad de las Fuerzas Armadas- ESPE, aiming to reach as many people as possible and reduce loss of life and damage [3].

The transition from analog to digital terrestrial television (DTT) in Ecuador, driven by the government since 2010, has been a catalyst for improving the dissemination of early warnings through this medium. Adopting the ISDB-Tb standard (International et al. - Terrestrial) has allowed for modernizing television broadcasts and incorporating additional services, such as emergency alerts [4].

The ESPE has played a crucial role in researching and developing technologies related to early warning in digital terrestrial television services. Previous projects, such as Software Defined Radio (SDR), were used to implement a terrestrial digital television transmitter and an emergency signal receiver. These have laid the groundwork for the work presented in this article [5, 6].

This research's significance lies in its ability to communicate early warnings to the population through multiple end devices, such as mobile phones, considering that a significant portion of the Ecuadorian population has access to these devices and the Internet. The rapid dissemination of information amid a natural disaster can save lives and mitigate the negative impact on society.

A project similar to the one presented in this article was developed in Peru by experts and manufacturers from Japan in collaboration with the Japan International Cooperation Agency (JICA) to enhance the transmission of emergency alerts. The development of the EWBS Gateway, which plays a vital role in the system, was explicitly requested by the National Peruvian Organization for Disaster Prevention. This device is designed to send emergency notifications to personal computers and mobile devices within internal networks, leveraging existing communication infrastructure to optimize the effectiveness and reach of the system in response to emergencies [7].

The present work focuses on implementing a Gateway Server that enables the reception and retransmission of messages from the early warning system in digital terrestrial television services. Mobile and desktop applications will also be developed to allow users to receive these alerts promptly. The project scope includes creating a private virtual server, a Python script, and applications for multiple platforms. Furthermore, future work will utilize technologies like Docker and Kubernetes to enhance the system's scalability.

This work aims to contribute to the safety and well-being of the population by providing an effective early warning system within the context of digital terrestrial television, offering the capability to communicate alerts to many people efficiently. Using cutting-edge technologies and collaboration with various institutions in Ecuador will enhance disaster response and help prevent more significant damages.

2 Methodology

To initiate the project, we developed an ISDB-T receiver with EWBS on SDR, building on the foundation in references [5] and [6]. Initially, we created a Python script to extract the overlay message from the transport stream transmitted via EWBS. Later, we used the message from the Radio on a software-defined radio (SDR) receiver and followed the methodology outlined below to advance the project.

Figure 1 depicts the project's comprehensive block diagram, representing a complete scenario of a terrestrial digital television transmitter and receiver, with its Transport Stream (TS) containing the early warning message. The "Gateway" section is particularly relevant to this research, which involves reading the message and retransmitting it through various mobile or desktop applications used by end clients. This aspect represents three stages of design and implementation.

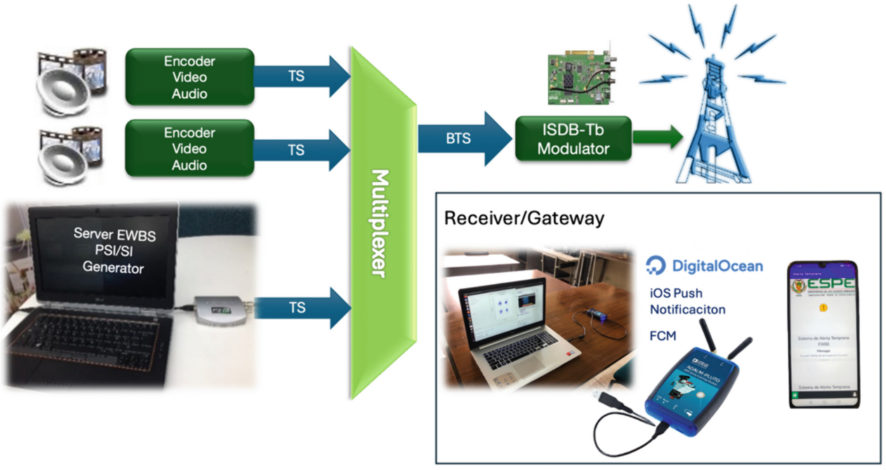


Fig. 1. Comprehensive Block Diagram of Terrestrial Digital Television System with EWBS and Gateway.

For the transmitter, we employed an ISDB-T modulator, specifically the DeKtec DTA-211, which, via the StreamXpress software, transmits the TS signal over VHF to UHF bands. This TS included an early warning message of the EWBS type, generated based on the project presented in [4]. However, any other TS conforming to the EWBS system under the ISDB-T standard could also be utilized.

The receiver, designed using GNU SDR platforms such as Adalm Pluto and based on the project outlined in [5], enables real-time capture of the OneSeg signal of ISDB-T, with the EWBS system configured within the Transport Stream (TS) structure and the design conducted in the UHF bands. We have created a Python script with multiple functions, such as reading early warning messages, triggering the alert signal, and notifying mobile clients. This script mediates between the received message and the mobile applications, ensuring that alerts are immediately disseminated. The script checks the file's content for emptiness and takes no action if the file is empty. However, when the transmitter issues an alarm with an encrypted message, the receiving computer via the SDR detects the alert, decodes the message, and saves the decoded message into a text file named "Alert.txt" at the path "/home/user/Downloads."

The Gateway Server component plays a crucial role in our project. It receives the TS from the receiver, efficiently separating the audio, video, and encrypted early warning message data. The EWBS Receiver block then receives and decodes the message, storing

it in a specific path. This data management process is a key factor in the smooth functioning of our system. Subsequently, a Public Server, VPS, or Droplet is implemented. This component handles data persistence and token management, serving as a query interface for desktop clients and ensuring the availability of necessary data, thereby facilitating communication between the server and desktop applications.

Our project was designed with a strong focus on the end client's experience. Web and mobile applications were developed to replicate the early warning message. These applications serve as the user interface through which end clients receive and interact with the alerts, ensuring prompt and effective dissemination of the message. This user-centric approach is a key aspect of our project, highlighting our commitment to effective communication.

EWBS Receiver

This section provides an overview of the processing blocks used for One-seg reception, as shown in Fig. 2 from the GNU Radio Companion environment, which were developed in the work for implementing a One Seg receiver [5]. The reception starts from the PlutoSDR Source block, which tunes and samples the signal from the Adam Pluto device at a specific rate allowed by the software, as indicated in the algorithm developed for the signal recorder since not all devices can sample at any rate.

The OFDM synchronization stage is intended to counteract timing and frequency synchronization effects in the OFDM system and is also responsible for orthogonal demodulation. The synchronization stage searches for the samples at the beginning of each symbol and estimates them to correct the carrier frequency offsets subsequently. The Sync Dem Ofdm 1Seg and OFDM Dem 1seg blocks perform this procedure. The first of these applies a Van de Beek maximum likelihood algorithm in a $2N + L$ window, where N represents the number of samples per symbol per operating mode (e.g., 1024 samples in a mode of operation 3), while L represents the length of samples in time for the guard interval [5]. Additionally, it determines the SNR (Signal Noise Ratio) corresponding to one of the objective metrics for evaluating performance and its variations in reception. It has an output of complex sample signals connected to the input of the OFDM Dem 1seg block.

This second block performs time and frequency synchronization by converting signals from the time domain to carriers in frequency (e.g., mode of operation 3 has 432 active carriers: auxiliary, TMCC, pilot, and data). It comes from the gr-isdbt module.

Subsequently, the most crucial block is the TMCC Decoder 1-seg, as it allows for the extraction of information concerning the activation of bit 26, where the emergency alert system of the TMCC carrier is located, which has 204 symbols. At its input, it has all the active carriers; at its output, only those corresponding to data.

Once the information is processed for one-seg reception, frames or packets of 188 bytes are outputted and received at the blocks' input (Read EWBS, Emergency Message, Flag EWBS). These processing blocks, programmed in Python for the GNU Radio environment, are responsible for extracting the characteristics of the EWBS emergency alert system. Figure 3 illustrates the processing blocks for EWBS. The Read EWBS block processes the data flow of 188 bytes (TS) at the receiver's output to extract detailed information from the PMT table and descriptor1 containing the emergency information.

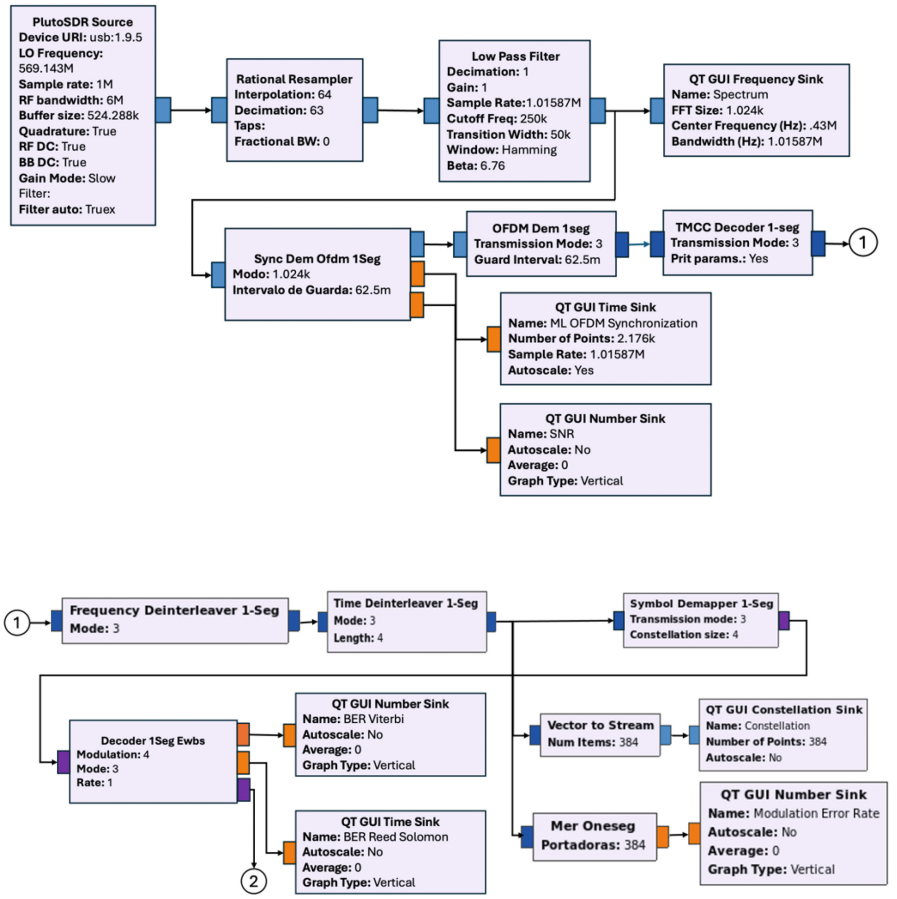


Fig. 2. GNU Radio Companion Environment for One Seg Receiver Implementation.

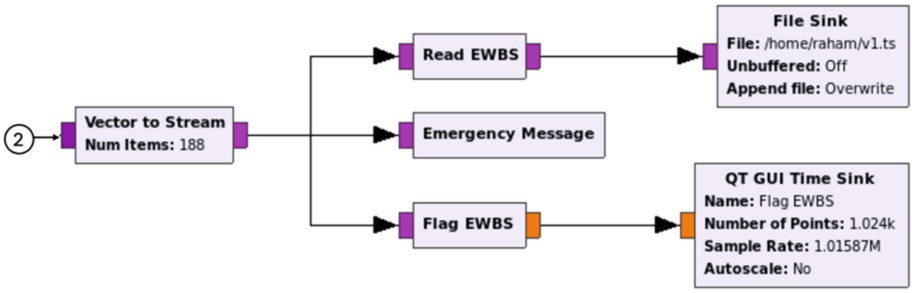


Fig. 3. Processing Blocks for EWBS.

Cloud Server Configuration
The initial stage involved selecting a cloud service provider, such as Digital Ocean [8], to host the server. Subsequently, a Droplet, a virtual machine on Digital Ocean, was created and configured with the Ubuntu 20.04 LTS operating system. A domain name was also associated with the Droplet's public IP address, necessitating DNS records configuration at the domain provider.

Dependency Installation
We have successfully installed all the necessary dependencies on the server, including MySQL 8 for data persistence, Python 3.8, Django, and the Django Rest Framework for creating and managing the web API. Additionally, we installed Nginx as a web server to act as a reverse proxy and Gunicorn to serve the Django application [9].

The database, an instance of MySQL 8 installed on the VPS, plays a crucial role in our system. It is used for data persistence, such as messages and tokens, and through its relational logic, it stores messages read by different clients.

Below are the different tables used in the MySQL relational database "Messages." Their usage will be explained later to ensure the proper functioning of the Gateway Server.

- **auth_user (Framework-specific)**: This table stores information about the superuser, who has permissions to save new messages and query tokens from different devices. An unregistered user cannot send alerts.
- **Messages**: This table stores early warning messages along with their respective timestamps or creation dates.
- **Tokens**: This table stores the token IDs (Device IDs) of each device that has installed the application, whether they are desktop clients or mobile apps. It contains two flags to identify if the token is mobile ('esMovil') and whether it belongs to iOS or Android ('esAndroid').
- **Token_messages**: This table establishes a many-to-many relationship between messages and tokens. By querying this table, we can determine which devices have successfully received the sent message.

Other tables within the DRF framework are also used for logging, sessions, and migrations, which could be more relevant to the project.

Django Application Configuration
This section involves cloning the Django project from a repository onto the server, setting up a Python virtual environment, and executing migrations to create the database tables. Additionally, a Django superuser was created to manage the application.

Django Rest Framework is a Python framework for web applications that optimizes code quantity when building web applications and handles application logic.

Application Deployment in Production
During deployment, the team configured a service to execute Gunicorn and established an Nginx configuration file for the domain. Furthermore, they enabled an SSL certificate through Let's Encrypt to ensure secure connections via HTTPS. Adjustments were made to the firewall to permit traffic on ports 80 and 443 [10]. On port 80, Nginx acted as a

reverse proxy, directing requests to Gunicorn, which can efficiently manage thousands of simultaneous connections, accessible at https://alertaewbs.site.

Gunicorn is an application server that translates HTTP requests into a language that Python can understand using the Web Server Gateway Interface (WSGI), a standard between web servers and web applications. Figure 4 illustrates how the architecture of Nginx, Gunicorn, and Python operates.

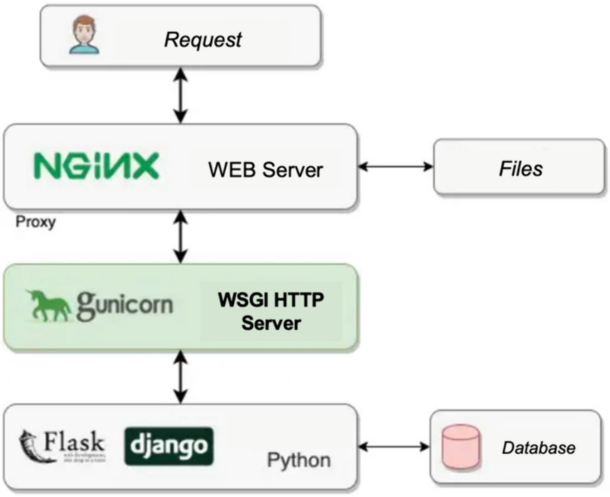


Fig. 4. Architecture of Nginx, Gunicorn, and Python.

Figure 5 shows a flow diagram of the receiver for reading an early warning message for subsequent dissemination on various platforms.

As a second security measure against potential attacks, two endpoints of the system were protected:

- /Insert/: Inserts a new message into the database.
- /Tokens/: Inserts and registers a new device (Device ID) in the database.

We protect the Droplet from potential attacks by using JSON Web Token (JWT) [11], an RFC 7519 standard that ensures secure data exchange between two parties and implements security measures commonly used in most systems and services. Authentication with Apple's APNs is performed in the same manner.

Development of Applications

This stage focused on developing desktop and mobile applications. The desktop application was developed using Spring Boot [12], while the mobile application was developed using React-Native [13] with Expo. Both applications are registered on the server and query unread messages. Push notifications for iOS and Android devices are implemented using APNs [14] and FCM [15].

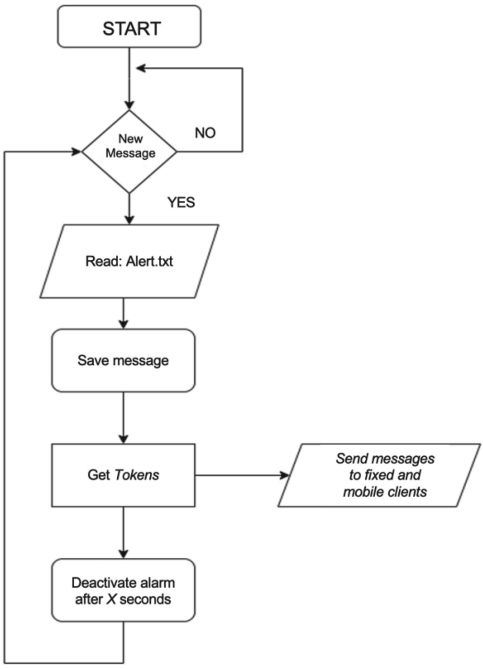


Fig. 5. Flow Diagram of Receiver for Early Warning Message Reading and Dissemination on Various Platforms.

Testing and Deployment
Exhaustive testing was conducted on all system parts to ensure its correct operation. Once all tests are successful, the system will be ready to enter production and provide early warning services to end users.

Maintenance and Updates
Finally, a continuous maintenance plan is established. The system is continuously monitored and maintained to ensure optimal operation. Software and hardware updates are applied as necessary.

3 Results

The tests conducted at the Digital Television Laboratory of the Universidad de las Fuerzas Armadas - ESPE yielded satisfactory results in the context of early warnings through digital terrestrial television. The devices involved in the test scenario included a transmitter utilizing a computer with StreamXpress software, a DeKtec DTA-211 modulator, and an antenna, alongside a receiver consisting of a computer and an Adalm Pluto SDR device. End clients were distributed across Windows, Linux, iOS, and Android systems, allowing for the evaluation of the system's applicability on various platforms.

Figures 6 and 7 document the testing carried out in a laboratory environment and demonstrate the system's successful operation. Notifications were effectively received

on both desktop and mobile devices, indicating the viability and robustness of the implemented infrastructure. The collaboration of various devices and components, such as the transmitter, modulator, antenna, receiver, computers, and the Software Defined Radio (SDR) Adalm Pluto [16], was essential in achieving this outcome.

Fig. 6. Video Demonstration in the Digital Television Laboratory of UFA - ESPE: Real-World Environment Showcase.

Fig. 7. Real-world Testing Environment at the Digital Television Laboratory of the Universidad de las Fuerzas Armadas – ESPE.

Java's JFrame notifications were utilized in the desktop environment, as depicted in Fig. 8. This enabled user alerts on both Windows and Linux operating systems.

For mobile devices, applications were developed for both Android and iOS systems, featuring intuitive and functional user interfaces. In the case of Android, two Tab Navigation layouts were created, offering a list of early warning messages and the ability to customize the alarm. Figure 9 displays the received notifications.

A similar approach was followed in the iOS environment, leveraging the versatility of cross-platform technology. Figure 10 showcases notifications on iOS devices.

Stress tests of the system and server were conducted using Apache JMeter, enabling evaluation of the server's performance under various workloads. The results obtained

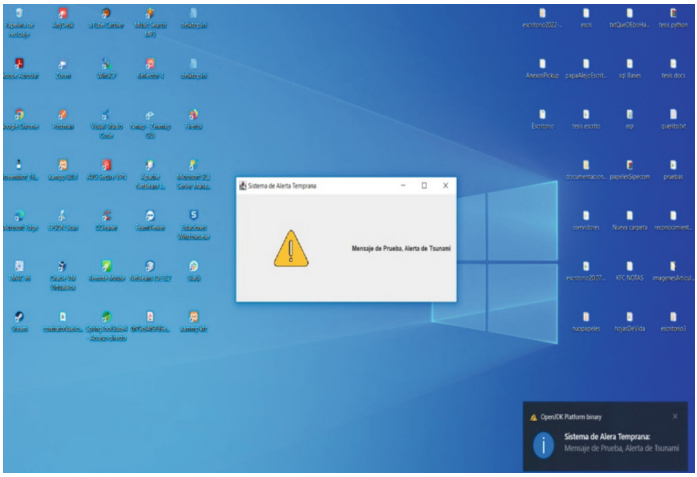


Fig. 8. Notification in a Windows 10 Desktop Environment.

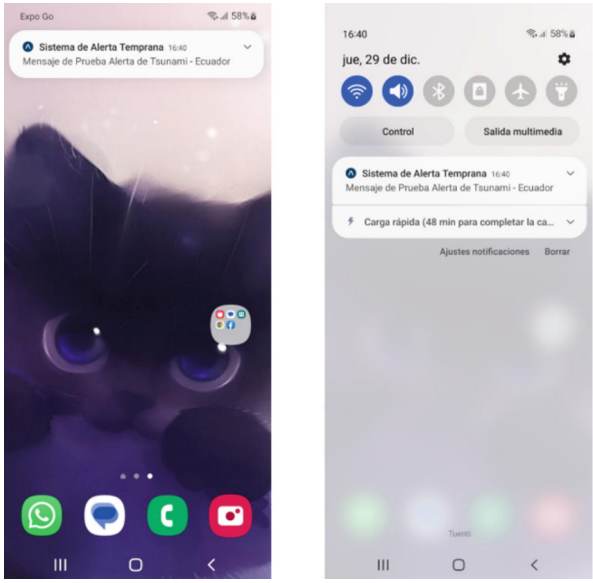


Fig. 9. Android Notifications.

are presented in Table 1. The server demonstrated excellent performance, even under significant loads, with minimal error rates.

Fig. 10. iOS Notifications.

Table 1. Stress Tests with Apache JMeter.

Number of Requests	Error Percentage [%]	Time [s]
100	0	1
200	0	3
300	0	4
400	0	5
500	3,2	7

4 Discussion

The implementation project of the Emergency Warning Broadcasting System (EWBS) in the context of terrestrial digital television represents a significant advancement in providing early warnings to the population. The integration of various components such as transmitters, modulators, antennas, and receivers has been demonstrated as crucial to ensure the effective transmission and reception of alerts in the terrestrial digital television environment. The collaboration of desktop and mobile devices allows notifications to reach a broad audience, increasing the likelihood that people will receive alerts in critical situations.

The server stress tests conducted with Apache JMeter have revealed the system's robustness and efficiency under increasing workloads. The results have shown minimal error rates even under significant loads, supporting the system's scalability and performance in a real-world scenario. Additionally, it was observed that the infrastructure used is highly effective in disseminating alerts on desktop and mobile devices.

5 Conclusions

In the context of terrestrial digital television, implementing the EWBS system has proven to be a promising technological solution for rapidly disseminating early warnings. The ability to notify through both desktop and mobile devices increases the accessibility of alerts, considering that a significant portion of the Ecuadorian population has access to these devices and the internet. The successful implementation of mobile applications for Android and iOS systems allows users to customize their notifications and access critical information promptly.

The server stress tests conducted with Apache JMeter have not just revealed, but strongly affirmed the system's robustness and efficiency under increasing workloads. The results have shown minimal error rates even under significant loads, providing a solid foundation for the system's scalability and performance in a real-world scenario. Additionally, it was observed that the infrastructure used is highly effective in disseminating alerts on desktop and mobile devices.

One of the significant changes in the Gateway design developed by Japan for Peru is the ability to upload area codes and emergency alert messages to the cloud. This structure allows for the extension of emergency alert information beyond the Gateway's coverage area and facilitates the integration of advanced options, such as displaying maps of the affected areas. Additionally, it incorporates the use of the Common Alerting Protocol (CAP). CAP is an international standard designed for the uniform transmission of alerts and warnings across various technologies, enabling effective propagation of emergency information to other devices. This scenario maximizes the existing digital terrestrial television infrastructure, ensuring broad dissemination and timely access to crucial information during emergencies.

References

1. Ecu911. Servicio Integrado de Seguridad, Sistema de Alerta Temprana. Recuperado de (2018). https://www.ecu911.gob.ec/sat-tsunami/
2. Takahashi, N., Nishida, K., Tsukada, S., Yamanaka, M., Horiuchi, S.: Earthquake early warning (EEW) in Japan: warning to the public and EEW research activities. Soil Dyn. Earthq. Eng. **29**(4), 798–810 (2009). https://doi.org/10.1016/j.soildyn.2008.06.008
3. Acosta, F., Sambrano, Y.: Diseño e Implementación de una Interfaz que Adapte una Señal de Emergencia de Televisión Digital a la red de telefonía IP de la Universidad de las Fuerzas Armadas - ESPE. Universidad de las Fuerzas Armadas - ESPE, Tesis de Ingeniería (2019)
4. Olmedo, G., Acosta, F., Haro, R., Villamarín, D., Benavides, N.: Broadcast Testing of emergency alert system for digital terrestrial television EWBS in Ecuador. In: Abásolo, M., Silva, T., González, N. (eds.) Applications and Usability of Interactive TV. jAUTI 2018. Communications in Computer and Information Science, vol. 1004. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23862-9_13
5. Olmedo, G., Castillo, R.: . Implementación de un sistema de recepción de una señal de emergencia EWBS utilizando la Plataforma de Radio Definida por Software para el estándar ISDB-Tb. Tesis de Ingeniería, Universidad de las Fuerzas Armadas - ESPE. Recuperado de (2021). http://repositorio.espe.edu.ec/bitstream/21000/25711/1/T-ESPE-044717.pdf
6. Olmedo, G., Chanataxi, B., Benavides, N.: Receiver for ISDB-Tb standard with low-cost SDR and interactive application extractor. In: 2022 IEEE ANDESCON, Barranquilla, Colombia, pp. 1-6 (2022).https://doi.org/10.1109/ANDESCON56260.2022.9989634

7. Sakaguchi, Y., Takahashi, Y., Sakuma, S.: Actividades de Difusión de la Tecnología Japonesa EWBS - Sistema de Alerta de Emergencias por Radiodifusión – (Emergency Warning Broadcast System)," presentado en DiBEG, Tokio, Japón. Disponible (2022). https://www.dibeg.org,
8. Matheus, C.: Rockcontent. Digital Ocean: qué es, cómo usar, ventajas y desventajas (22 de abril de 2020). Recuperado de https://rockcontent.com/es/blog/digital-ocean/
9. FIB, i. (7 de julio de 2016). Django API REST. Recuperado de https://inlab.fib.upc.edu/es/blog/django-api-rest#:~:text=Django%20REST%20Framework%20es%20un,o%20el%20proyecto%20UOC%20%C3%ADndex
10. Gustavo, B.: Hostinger. ¿Qué es SSL, TLS y HTTPS? Recuperado de (22 de noviembre de 2022). https://www.hostinger.es/tutoriales/ssl-tls-https
11. López Magaña, I.: Qué es JSON Web Token y cómo funciona. OpenWebinars (2020). Recuperado de https://openwebinars.net/blog/que-es-json-web-token-y-como-funciona
12. Wigglesworth, A., Duchi, J.: Pro Spring Boot. Springer (2016).https://doi.org/10.1007/978-1-4842-1890-0
13. Facebook, Inc. React Native. GitHub (2013). https://github.com/facebook/react-native
14. Apple Inc. Local and Remote Notification Programming Guide. Apple Developer (2021). https://developer.apple.com/library/archive/documentation/NetworkingInternet/Conceptual/RemoteNotificationsPG/
15. Google LLC. Firebase Cloud Messaging. Firebase (2021). https://firebase.google.com/docs/cloud-messaging
16. Analog Devices. Adalm Pluto SDR Active Learning Modulo. Recuperado de (2017). https://www.analog.com/media/en/news-marketing-collateral/product-highlight/ADALM-PLUTO-Product-Highlight.pdf

Personalized Enrichment of Video Through a Crossmedia Environment

Alcina Prata[1(✉)], Teresa Chambel[2], and Miguel Sales Dias[3]

[1] Superior School of Technology (ESTS), Polytechnic Institute of Setúbal, Setúbal, Portugal
alcina.prata@estsetubal.ips.pt
[2] Lasige Faculty of Sciences, University of Lisbon, Lisbon, Portugal
mtchambel@ciencias.ulisboa.pt
[3] ISTAR_ISCTE, Instituto Universitário de Lisboa (ISCTE-IUL), 1649-026 Lisbon, Portugal
miguel.dias@iscte-iul.pt

Abstract. Crossmedia systems have been gaining space due to their flexibility and capability to deal with a diversity of contexts. The success of these systems resides, mainly, in their capacity to provide flexible solutions to the user interaction through different devices, especially if each device is used to perform what it is most suited for, from the user experience perspective. This paper briefly addresses the challenges faced when designing Crossmedia systems capable of generating personalized information that is added to the video. To illustrate our research, we designed a system, which evolved from previous versions, referred to as eiTV, which generates personalized web-based content that provides additional information on users' selected topics of interest, while watching a specific news video, such as a documentary or a series. The web-based content may be generated, accessed, personalized, shared and (immediately or later) viewed through iTV, PC or mobile devices, depending on the users' needs. The paper main contribution is the interaction model designed to support the 'create' functionality (which allows to choose the topics of interest) in the news videos, supporting iTV, PC and mobile. We created high-fidelity prototypes of the eiTV system, which were evaluated with the participation of 30 elements, with age range ranging from 20 to 55 years old. The achieved results are presented and discussed.

Keywords: Crossmedia · Transmedia · Video · News Videos · Web-based content · iTV

1 Introduction

Since its appearance in the nineties, crossmedia and transmedia systems have become a reality [1, 2]. These systems characteristics, the number of involved devices and their capability to support a diversity of contexts of use through flexible and personalized solutions, were some of the reasons for their success in a multiplicity of areas [1–3]. The success and adoption of these environments can be explained by different factors, like the use of new and appealing devices, social networks, technological advances, user changes in terms of interests and habits, and specific systems characteristics, which the most relevant are, flexibility, diversity, and mobility, so essential to support today's lifestyle [3] and a world that, in fact, is crossmedia in its essence [4].

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 31–48, 2024.
https://doi.org/10.1007/978-3-031-70439-0_3

One area that highly benefited from crossmedia systems was learning [5–7], independently of being formal or informal. As to the medium used to support learning through crossmedia systems, video is one of the richest and favorite ones, while the preferred devices to access video (even when referring to TV consumption), are TV, PC, and mobile devices [8, 9]. There is no universal preferred device, and the preferences may vary and depend highly on many factors like age range, technological literacy, mobility needs, consumption habits, cognitive condition, amongst others, and each device may open the door to flexible environments through structure and interaction. However, the design of these crossmedia environments/systems faces some challenges that, may affect their effective use and adoption, and need to be addressed [10, 11].

To illustrate our research, we designed and developed the eiTV system. The name was chosen to suggest the use of the Internet to provide additional content to the interactive TV, not through search but as part of a crossmedia system, in a complementary, structured, personalized, and interconnected way. The eiTV system has been through an evolution process of 4 generations of prototypes, ranging from low to high fidelity prototypes. The first-generation conceptual goal was to explore the design of an application capable of generating, from iTV, personalized web-based content as additional information to the program being watched, in response to informal learning opportunities, to be seen through PC, TV or mobile. The main concern - here and along all the other generations - was to explore the model and functionalities that better supported viewers change in cognition modes (also implying contributions to the application flexibility and personalization), continuity across devices, contextualization, and User Experience (UX), optimizing the use of each device [12]. The second-generation conceptual shift was based on a 'beyond iTV' desire as well as with the design of a portal instead of an isolated application. Thus, we may say that this generation is more aligned with the concept of 'going beyond iTV in the CLOUD. The paradigm changed due to technological and social factors. Video can be watched anytime, anywhere, from different types of devices (not only iTV as in the first generation). Each device (TV, PC and mobile) may be used to watch video, create the associated web-based content, and access it. This allowed a natural evolution to a broader video-based application and an evolution to a portal with more refined functionalities, a relevant evolution considering that viewers no longer want to be passive. They want a more active role, to collaborate, and to create [13]. In this context, it is acceptable to say that consumers have turned into active producers, a role that becomes a true possibility within a portal with these functionalities. The third-generation conceptual focus was on mobility. Conceptually, the keyword here is MOBILE, 'going mobile', and the flexibility inherent of being mobile with the co-existence of different devices and contexts of use. The goal was to take the best advantage of smartphones, in terms of mobility and specific features, as for instance: use the mobile GPS to access content through its geo-location and contribute to the enrichment of the application with geo-referenced content. Another goal was to take advantage from their (complementarity) synchronization with other devices, to simultaneously show different but related information on different devices. This is usually referred to as the 'second screen' phenomenon. As an example, watching the video on the computer while using the mobile device to watch the generated web-based content about that video, contributes to flexibility, personalization, and adaptation to different

cognitive modes [14, 15]. Finaly, the fourth-generation conceptual foundation keywords were DIVERSITY, 'diverse types of videos', and SOCIALIZATION, 'support viewers social communication needs while watching the videos', two research opportunities identified in our previous work. The first three generations allowed us to achieve a stabilized model for the design of crossmedia environments able to create, from the videos being watched (on iTV, PC, or mobile), additional web-based extra information based on the viewers topics of interest, chosen while watching the video. However, all the previous prototypes were developed based on the series videos genre and, in nature, series videos are less dynamic than videos about documentaries or news, which are typically more dynamic, imply a different cognitive mode, and are watched in a lean forward mode, thus creating new opportunities for research. So, the fourth generation's main concern was to adapt the existent interaction model to conceptually accommodate diverse types of videos, providing the system with a different level of flexibility and continuity. We started by researching the dynamics and characteristics of documentary videos and then, we redesigned our system interaction model to meet our findings [4, 16, 17]. Next, we conducted the study presented in this paper, whose main goal and contribution was the redesign of the interaction model and interfaces to accommodate news videos without conflicting with the previously conceptualized eiTV solutions to accommodate series and documentary videos. Concluding, our research question is: *is it possible to create an easy, intuitive, natural, and non-intrusive interaction model that allows interacting with news videos in the context of the previously developed eiTV system in an integrated manner*?

As to the structure of this paper, Sect. 2 presents a review of related work and concepts. Section 3, presents a description of the design challenges associated to crossmedia applications. Section 4 describes some of the most important design decisions, while Sect. 5 describes the evaluation process. Finally, Sect. 6 presents the conclusions and perspectives for future research and developments.

2 Related Work

The TAMALLE project [18] developed a 'dual device system' for informal English language learning, based on watching iTV and selecting what to access later, on the mobile phone. This was an interesting system capable to accommodate different cognitive modes and different contexts of use, especially, if considering the mobile phone possibilities. This work was important to our research due to the good results achieved by providing users with mobility in the use of the system. Obrist et al. [19] developed a "6 key navigation model" and its interface for an electronic program guide running on the TV, PC, and mobile phone. The different devices were not used in a complementary way since the intention was to test a similar interface, on three different devices. They have perceived that viewers prefer fewer navigation keys and a unified UI with the same functionalities across devices. This confirmed our prototypes UI design last decisions. Newstream [20] provides extra information about what is being watched and related websites, using TV, PC, and mobiles. Depending on the viewers' needs, that extra information may be viewed immediately, stored for later view, or pushed to another device. Each device maintains awareness of each other and is able to: move interaction to the

device that makes the most sense in a specific context, use several devices simultaneously, and use the mobile device as a remote to the TV and PC. Limitations include: the system relies almost exclusively on social networks to receive and share content, for interaction and dialogues; and the limited viewer direct influence on the new content presented as extra information. Our work is more flexible in these concerns. 2BEON [21], currently called WeOnTV, is an iTV application which supports the communication between viewers, textually and in real-time, while watching a specific program. It also allows viewers to see which of their contacts are online, which programs they are watching, and instant messaging on the iTV, demonstrated to be important to give viewers a sense of presence and was implemented with smartphones as "secondary input devices". This work demonstrates the importance of sharing information with viewers' contacts about what they are watching on TV, which supports our own decision to include a sharing functionality in eiTV. Cronkite [22] provides extra information to viewers of broadcast news. While viewers are watching a news story, they feel the need to know more about it, they press the "interest" button on their remote and the system provides them with extra information on the computer display. The extra information is about the story that they are watching rather than specific topics of interest inside the story, which is somehow limited. To have the system working, both TV and PC need to be simultaneously on. The system is limited considering that the extra information is not stored for the latter view (and that might be the viewers' preference). Our application stores the related information for later use and the simultaneous use of iTV and PC is a possibility but not the only option available. In fact, viewers may select very specific topics of interest inside a story instead of the whole story and some specific functionalities, as asynchronous communication tools, were also contemplated.

In a recent study about operation latency in the context of crossmedia systems [11], something that was conducted for the first time, the authors explored the effect of operation latency on user experience and operation performance for four cross-device tasks. They found that, as the operation latency increased, user satisfaction dropped, and they presented specific measures (in terms of time) that will help us in our future prototypes' evaluation process.

3 CrossMedia Design Challenges

This section briefly describes the main cognitive and affective aspects that need to be considered to effectively design crossmedia environments and interfaces, with a special focus on the design challenges associated with video and different devices. Considering that all our prototypes were developed based on these fundamental aspects, they were briefly included here for contextualization purposes. Their detailed explanation may be found in our previous work [12].

Media and Cognition: Norman's view [23] defines two fundamental cognitive modes: experiential and reflective. The experiential mode allows us to perceive and react to events naturally and without cognition, but requires different technological support, and the medium affects the way we interpret and use the message and its impact on us. To exemplify, TV and video are typically watched in an experiential mode while learning strongly relies on reflection. A successful integration of media should have into account

what each medium and device is most suited for in each context of use, augmenting and complementing their capabilities in a flexible combination.

Crossmedia Interaction, Conceptual Model and User Experience: The main challenges of crossmedia interaction design described by [24] include: consistency, interoperability, and technological literacy needed for the different devices. The conceptual model, how the software will look like and act, is also a very important aspect since several interaction scenarios and contexts are involved [26]. The quality of the interaction cannot be measured only by the quality of the system parts, but as a whole. In this context, the user experience (UX) may be evaluated through how well it supports the synergic use of each medium and the different kinds of affordances involved, also understanding what makes the user pass the current medium boundaries to use other media as well. According to [26], the UX may involve the isolated perception of the medium (distributed), one of the biggest barriers to its efficient use and adoption, or the perception of the system as a whole unity (coherent). According to [19], the UX evaluation methods and measures relevant, when ubiquitous TV is involved, are: physiological data; data mining, log files, observation, case studies, lab experiments, experience sampling method, probes, diaries, interviews, surveys and focus groups. The combination of methods to use depends on each specific case.

Supporting Crossmedia HCI: In this context, the migration of tasks is supported via crossmedia usability and continuity, influencing on how well and smoothly users' skills and experiences are transferred across the different devices [27] and contexts of use. The consistent look and feel across media is an important requirement, even if it should not limit the goal of having each medium doing what it is most suited for and extending its characteristics (synergic use) [28].

Designing for Different Devices and Contexts of Use: Crossmedia design involves designing interfaces for different devices. To understand the devices and have each one performing what it is most suited for, the best approach is usually to study each situation, including device characteristics and cognitive and affective aspects associated with its use: why people use them, in which mode, compare them, etc., and the design guidelines for each device [5] followed by an adequate combination.

Supporting Communication Tools in Crossmedia Contexts: The use of communication tools integrated with crossmedia contexts and environments requires the understanding of each device and media characteristics. Each device should contribute with what is most suited for while the media should not change its nature by the incrementation of new tools. Every type of communication tool, in this context, should be designed mainly with UX in mind. To minimize the complexity associated to this type of contexts and take the best advantage on the users' previous knowledge, the Interface should be as close as possible from the more traditional use given to that kind of tool (if an already existent one is being replicated in a different device) [4].

4 Crossmedia Design in EiTV

This Section briefly presents the main functionalities and the design options concerning the eiTV Crossmedia system, in response to the design challenges identified in Sect. 3. Thus, the different contexts of use and the devices characteristics and cognitive and affective aspects associated with their use, were carefully considered in the design process. Specific attention was given to the 'Create' functionality considering that it refers to functionality interaction models proposed and tested in this paper. As to the technical details on the system's backend, such as algorithms or data processing methods, they are outside the scope of this work and thus were not included.

4.1 Architecture of the EiTV

In terms of architecture, the eiTV system works as a portal aggregator of all the functionalities. The portal may be accessed from any of the devices (iTV, PC and mobile devices), thus working as a true 'ecosystem of devices'. Through the portal it is possible to: generate web-based content; see, edit, and share web-based content, upload files, change profile, communicate via chat, etc. Concluding, everyone may receive web-based content generated by the eiTV, a characteristic that provides flexibility to the application.

4.2 Navigation Model Based on Flexibility

As to the navigation model, a menu style navigation was implemented. This solution provides users much more control over their choices considering that all the functionalities may be accessed at any moment, directly through the menu or through the chromatic keys (when using a remote). This navigation model improves: the application interoperability since it shows people how it works; the UX which becomes more coherent considering that users easily perceive the system as a whole unit; the crossmedia interaction continuity through different devices and the interaction consistency considering that it becomes easier to reuse users' interaction knowledge. Due to its flexibility, this model is also more adapted to changes in cognition modes, levels of attention and technological literacy.

4.3 Functionalities of the EiTV

a) The **Create functionality** allows users to watch videos and select topics of interest for further information. On this study prototypes, the topics of the used videos are about news. Due to the dynamic nature and characteristics of these videos, several topics of interest may arise almost simultaneously and should be made available in an integrated, contextualized, flexible and not intrusive way. Four interaction models were designed and tested in low fidelity prototypes which, after evaluation, allowed us to choose the two final interaction models (a1 and a2, which can be seen in Fig. 1). Both a1 and a2 were designed and tested in high fidelity prototypes, and work as explained next:

a1) In the interaction model a1) five selectable topics appear onscreen, in white color, and with the most likely selectable topic in higher font size, underlined and bold.

The topic underlined is always the one in the third position and corresponds to the one being talked about at that moment. So basically, the first 2 topics correspond to issues that were already talked about, the third topic is the one being talked about at that moment, and the fourth and fifth topics are the next on the line and they will move to the third position when being talked about. Regarding the underline, it should be noted that the underline bar doesn't appear immediately but progressively from left to right. The purpose of this designed option was, mainly, to avoid UX conflicts since, when outside iTV, the underline is typically used as a link instead of a method to highlight a specific word. Another goal was, to provide some dynamics and visual feedback to the selectable topic, in a way that could be easily perceived and understood, without disturbing the video viewing experience. The list of topics appears on the up-left corner of the screen, over a transparent curtain, and moves automatically as a scroll down menu while the narrative goes on (see Fig. 1, interaction model a1). To choose a topic, the user just needs to press the OK button (or use the mouse or touch screen - depending on the device being used) while the topic is active. The availability of 5 topics simultaneously helps create a sense of continuity and provides the users with more time to decide if they want to choose the topic or not thus providing more flexibility in a very dynamic context. As to the use of a scroll down menu style, it activates the UX by recreating a type of interface commonly used online. To provide visual feedback on the choices that were made, the selected topics appear with the ✓; symbol in front. Important to mention that this option is activated by default but, to assure personalization and a low level of interference, it may be easily deactivated by pressing the "i" button.

a2) In this interaction model the list of 5 topics appears on the same place and with the same dynamics. The difference is that, instead of underlined, the selectable topic appears with a blue 'shadow bar' while active (see Fig. 1, interaction model a2). Using a blue shade that appears while the topic is being pronounced activates the UX by recreating other interfaces commonly used (as for instance the Orwell game) that are perceived as usable and intuitive due to the feedback that it provides. By using a design solution that is typically used through the web in an iTV interface, we expected to create a certain continuity amongst the different devices and experiences as well as provide the system with more flexibility.

In both interaction models a1) and a2), the information about a topic was made available in three levels, from less to high informative (as presented in Table 1):

At any moment, the user may change between levels of information by pressing buttons 1, 2 or 3 or by using the directional buttons or by using the mouse or touch screen (depending on the device being used). Thus, the eiTV navigation is adaptable to users with different technological literacy. It was decided to maintain the 3 levels of information, with embedded and overlaid options on levels 2 and 3, since we saw from previous prototypes [13, 14], that they play an important role to accommodate users' changes in cognition modes, levels of attention, goals, needs and interaction preferences. These options also give users the possibility to personalize their viewing/interactive model, so important when video is involved [4].

Table 1. Levels of information about a topic

Level	Description/Scope	Interface:
Level 1 (Topics)	Only implies the use of the designed solutions to select topics of interest	1. Remote OK button when watching the video from the TV 2. To touch the screen (when watching the video from the mobile) 3. Use the mouse or touch the screen when watching from the PC)
Level 2 (Summary)	Implies the immediate display of extra information as a brief summary about the topics	The immediate extra information appears: 1. Overlaid onscreen or 2. Embedded onscreen
Level 3 (Structured)	Implies the immediate display of extra information, namely a structured list of that topic main aspects or options that the user may choose	The immediate extra information appears: 1. Overlaid onscreen or 2. Embedded onscreen

For contextualization purposes, Fig. 2 presents the interaction models designed and tested for different video genres: CSI series [previous work published in 14, 15], documentaries [previous work published in 16, 17] and news videos (the contribution of this paper). Due to space constraints, only the interfaces for iTV are presented, but they were also designed and tested for PC and mobile.

b) Search functionality allows searching videos based on different criteria, namely, video criteria and system criteria. The Video criteria refers to a search made by title, actor name, etc.; and the system criteria refers to search video with or without web content(s) already generated. These two options provide the system with flexibility.
c) The Share functionality is available only after users accessed the Create or Search functionalities for the first time. The share functionality allows sharing the generated web-based content or the retrieved video (with or without web-based content), with their contacts thus providing users with flexibility. On this functionality personalization and error prevention were provided.
d) The User Profile functionality allows users to create their profile, manually or by importing their personal data from social networks (under their strict consent), allows edition, validation, etc. The user profile information is used to personalize the web-based content, thus improving flexibility.
e) The DF functionality was designed to have each device doing what it is most suited for and thus take the best advantage on each device's characteristics. Thus, the number of available functionalities depend on the device being used. In the case of mobile devices functionalities, the following were made available:
e1) Great flexibility and mobility (use it everywhere, anytime, anyway): When using the TV, the scroll is not an option, but that does not happen when using the other

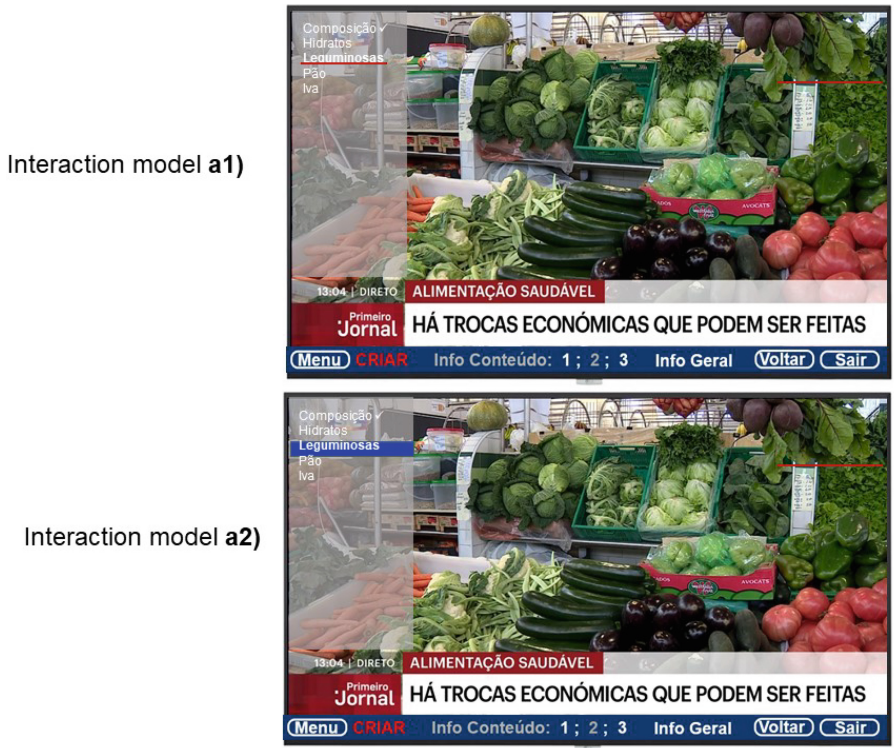


Fig. 1. Proposed interaction models (in this case the iTV is being used and we are in the information level 1, the less intrusive considering that no extra information is made available)

devices; contrary to TV, small PCs provide some mobility, and mobile devices may be used everywhere, even when users are standing up, meaning that any extra time may be used to interact with the system;

e2) Location-based search using the GPS functionality: The search functionality allows users to search videos related to their current location. As an example, when near the liberty statue the user may use this functionality to search, from its own system and the internet, videos related to that specific spot (this type of video files need to be inserted when using iTV or PC);

e3) Add immediately, or latter, shot pictures or videos, that may be related to the video being watched, as additional information to the web-based content or, instead, really integrated as part of the web-based content.

f) The Devices Synchronization functionality gives users the possibility to synchronize devices and make them work as a true ecosystem of devices. This provides the system with flexibility, continuity, and unity.

g) The Chat functionality was designed and made available to provide users with more flexibility and to support their social communication needs (so important amongst the youngest population) without losing the sense of unity (considering that they don't need to use external means and devices to communicate, just the ones included in the eiTV system in a specific an integrated way).

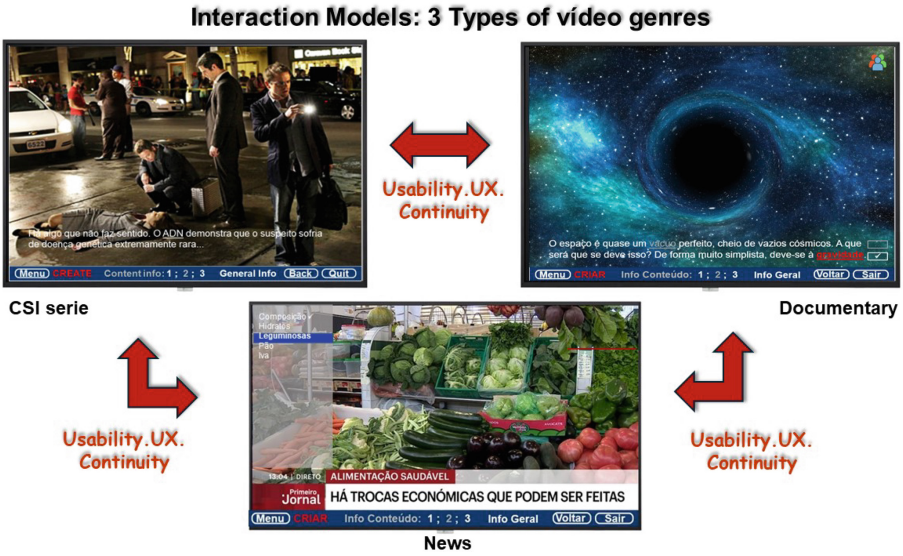


Fig. 2. Interaction models for the three types of videos: series, documentaries, and news

Consistency in UX and the perception of the system coherent unity independently of the device being used was also a priority. We decided to keep a coherent layout in terms of colors, symbols and other graphic elements, as navigational buttons, to better contextualize users, give them a sense of unity in their UX and to allow a smooth transition amongst media and devices. This allowed to provide users with a sense of sequence and continuity, respect the context of use and be consistent in terms of look and feel and navigational options in all the devices, and to help the perception of the application as a unity. Users are aware that they may access their eiTV system through different devices whenever they create web-based content, helping to conceptually understand the system as an 'ecosystem of devices'.

Important to mention that, for the purpose of this paper, only the create functionality was implemented in high fidelity through FIGMA software.

5 Evaluation

As mentioned, the formulated research question was: *is it possible to create an easy, intuitive, natural, and non-intrusive interaction model that allows interacting with news videos in the context of the previous developed eiTV system in an integrated manner?* From this expression, it was possible to conclude that we needed, at least, to evaluate two things:

1) The proposed interaction models for the news videos that were developed in this work (considering that they had to be usable for iTV, mobile and PC);
2) The integration of the most suitable one (in this context) with the previous developed eiTV system (which resulted from our previous work) where documentary and SCI series videos were tested and were running from iTV, mobile and PC.

In any type of system and context, UX evaluation is very important [30]. The UX evaluation methods and measures considered relevant for this specific case were: observation (internal and external), case studies, lab experiments, questionnaires (USE, NASA TLX, usability heuristics) and interviews.

Important to remind that, in a preliminary phase, four 'Create functionality interaction models' were tested in low-fidelity prototypes. That evaluation (with 4 usability experts and 15 end-users), allowed us to choose the two interaction models (a1 and a2) that were implemented in the high-fidelity prototypes presented and discussed in this paper. Considering that it was a preliminary phase, and due to space constraints, those four interaction models, and their evaluation, were not presented here.

The evaluation of the final high-fidelity 'Create functionality interaction models' a1) and a2) and the evaluation of the chosen one (a2) integrated with the existent eiTV system, occurred from March to October 2023, with the participation of 30 end-users (ranging from 20 to 55 years old) which were grouped into 2 evaluation groups: Group 1 (G1) composed of 15 participants, aged between 20 and 55 who already participated on previous evaluations and Group 2 (G2) composed of 15 participants, aged between 21 and 51 who were never in contact with the eiTV system. Inside each group, the participants were categorized into 3 subgroups as follows: 5 with high technological literacy; 5 with medium technological literacy and 5 with poor technological literacy. All the subgroups included 1 person below 25 years old, 2 persons between 25 and 40, and 2 persons above 40. As to the participants technological literacy categorization, it was possible via the use of a questionnaire with questions as: do you use Internet? e-mail? Facebook? Instagram? WhatsApp? How many hours a day? From which devices? etc. The idea of using a group of users that already participated on previous evaluations (the ones from G1) was to understand to what extent this system resulted more complex and/or hard to use when compared with the previous versions. As to using a group of evaluators that never interacted with the system (the ones from G2) the intention was trying to perceive how easy, useful, interesting it was for them, how usable the interfaces were and, amongst other factors, what impact the application had on them, particularly, considering this new level of complexity. Is important to refer that the described evaluation methodology was used with previous versions of the eiTV system with quite success [12–16] and, that was why, we decided to replicate it.

The evaluation process started with a demonstration of the high-fidelity prototype using the three devices involved (iTV, PC and mobile) and the functionality being tested. Next, the evaluation process was carried out in 3 major phases as explained:

1) **Phase 1** – users were asked to perform tasks that allowed us to evaluate the specific interfaces (a1 and a2), designed for choosing topics of interest, from the three devices, while watching news videos. Each user started by visualizing a **news video** and, after choosing topics of interest (from interaction model a1 and using the 3 information levels – described in Table 1) they generated a web-based content as additional personalized information to that video. Next, they repeated the process using interaction model a2. These activities were performed first through the iTV, next through the PC, and finally through the mobile. Finally, users were asked to fill a questionnaire and were interviewed. The questionnaire, used in all the 3 major phases, was based

on the USE questionnaire (usefulness, satisfaction, and ease of use) [31]; the NASA TLX questionnaire (cognitive overload) [32]; and usability heuristics.

We also wanted to evaluate the interaction model integrated with the existent eiTV system, so we needed to understand where the eiTV designed system was perceived, amongst other things, as more useful, and easy to use (if when interacting with news, documentary, or series videos). Thus, we proceeded to the next phases:

2) **Phase 2** - users were asked to view a **documentary video**, choose topics of interest, and generate a web-based content through the iTV. Then, users were asked to fill the questionnaire and next they were interviewed;
3) **Phase 3** – the process described in phase 2 was replicated here but users were asked to view a **series video** (from the CSI police series), instead of a documentary video.

Considering that on phases 2) and 3) our main priority was the evaluation of the iTV interaction models all the tests were performed in one contextual scenario (a simulated 'living room').

As to the interaction models proposed (a1 and a2), and in general, both models were easily used, adopted, and considered easy and natural to use but, as may be seen from Table 2, the preferences changed according to the device being used and the group. When using iTV, model a1) was the preferred in G2 (60%) while model a2) was the preferred in G1 (53%). No one found the models too much intrusive or confusing and the majority of the G1 elements that preferred a2) were the ones that participated in the first fourth generation prototypes meaning that, for them, the blue bar used to provide users with visual feedback was familiar. It was interesting to see that when there was no previous memory of the system (users from group G2) the chosen model was a1), the one without the blue bar. This seems to indicate that, when there is no previous knowledge of the interface, the preferred interaction model is the one that mostly resembles the interaction model of some popular TV channels. This makes sense considering that more than 50% of the users that have preferred the interaction model a1) are used to interact with the iTV in a daily basis. When through PC and mobile, a2 was clearly the preferred interaction model in both groups. When asked why, users with more literacy stated that they felt this interface is more visible, natural, and similar, to what they are used in other contexts. This was expected considering that 'Shadow bars' are commonly used in terms of interfaces to highlight text everywhere (online and offline) and in different environments and contexts of use. As to interaction model a1), and when from PC and mobile, 3 users stated that, in the beginning of the interaction, the use of underline was misleading since they thought it was a link, instead of a highlighting tool.

Considering that each device has its own characteristics, contexts of use and UX, a crossmedia environment represents a big UI challenge in terms of design. It is mandatory to assure continuity and contextualization with a smooth transition across devices, but this doesn't mean to use the same UI across devices. Thus, it is essential to use the most adequate solutions for each device without losing the feeling of continuity.

Considering the good results achieved and that the differences between interaction models were statistically significant, 2 solutions may be used: 1) use interaction model a1 when from iTV (the preferred in G2) and use interaction model a2 when from the other devices or; 2) use interaction model a2 in all the devices, considering that it was

Table 2. Preferred Interaction Model in each device

	eiTV		PC		Mobile	
Interaction Model	G1	G2	G1	G2	G1	G2
a1)	6 (40%)	**9 (60%)**	4 (27%)	5 (33%)	6 (40%)	5 (33%)
a2)	**8 (53%)**	6 (40%)	**11 (73%)**	**10 (67%)**	**7 (47%)**	**8 (53%)**
none	1 (7%)	0 (0%)	0 (0%)	0 (0%)	2 (13%)	2 (13%)

considered intuitive and well-designed, even for the 60% of users that preferred the a1 model when from iTV.

Yet about the interaction models proposed, both medium and high technological literacy categories reacted well to difficulties. In both (G1 and G2) low technological literacy subgroups, as expected, the older participants reacted with more resistance and discouragement in the presence of difficulties than the users below 25. In fact, 2 users from G2 (one with 41 and one with 48 years old) felt some difficulties while interacting with the news videos from iTV. In spite being more engaged in the use of iTV than the youngest generations, the older generations are not so used to digital interaction and, when they are, that tend to be associated with mobile phone interaction models.

Concerning the preferred device to generate the web-based content, from Table 3 it is possible to see that group G1 preferred the mobile (47%), while G2 preferred the PC (40%). As usually, iTV was the less preferred device (in both groups). However, in general, the iTV interaction models were considered useful (26–87%), easy to use (24–80%), intuitive (25–83%) and not intrusive (23–77%). What was important to perceive was that the iTV being the less preferred device happened due to mobility and consumption habits and not due to bad interfaces.

Contrary to what happened before, in group G1 the preferred device to create the web-based content was the mobile device. That happened amongst those that are used to watch TV from the mobile device. When considering both groups, the majority found the mobile phone interaction models useful (28–93%), easy to use (26–87%), intuitive (27–90%) and not intrusive (25–83%), contrary to what happened before, the size of the screen was not an issue when it comes to choose topics of interest. The results indicate that users didn't feel any constraint related with the mobile device characteristics.

Table 3. Preferred Device to Generate the Web-based Content

Device to generate the Web Content	G1	G2
iTV	3 (20%)	4 (27%)
PC	5 (33%)	**6 (40%)**
Mobile	**7 (47%)**	5 (33%)

In terms of the 'create functionality' information levels, and when considering the selection of topics from the news video, as can be seen from Table 4, the preferred

information levels were 1 (when through PC and iTV) and 2 (when through mobile), independently of the group (G1 or G2). This was somehow expected considering that these two levels are the less intrusive. However, it was interesting to see that, when through the mobile, the preferred information level was number 2. Considering the small size of the screen this was not expected. However, we noticed that, from the 8 G1 users that chose level 2, seven were below 35 years old. In fact, the new generations are more used to actively interact with the mobile device and use complex applications, which probably explains their preference for information level 2. In spite not being the preferred to anyone, information level 3 was considered very useful, mainly when from the PC and older population. These results clearly indicate the importance of keeping the three information levels, even when from news videos, which have a specific dynamic nature.

Table 4. Preferred Levels of Information through each device

Level of information when choosing topics of interest/device used	iTV		PC		Mobile	
	G1	G2	G1	G2	G1	G2
Level 1	**7 (47%)**	**9 (60%)**	**7 (47%)**	**9 (60%)**	6 (40%)	7 (47%)
Level 2	**7 (47%)**	6 (40%)	5 (33%)	5 (33%)	**8 (53%)**	**8 (53%)**
Level 3	1 (7%)	0 (0%)	3 (20%)	1 (7%)	1 (7%)	0 (0%)

As may be seen from Table 5, the possibility to choose topics of interest to know more about the topics that are being discussed, was considered more useful in documentary videos (50%). In spite acknowledging documentaries as the type of videos that mostly benefit from this type of system, users have used it more while watching news videos (43%). They've used it less while watching videos about the CSI series (50%) which was expected due to these videos' low dynamic nature. As to the system ease of use, 63% stated that CSI series was the easiest what makes sense considering that video series is the most simplified interaction model. Users found documentary the hardest type of video (43%) to choose topics of interest from. This was somehow unexpected considering that the news video genre its more dynamic and thus imply a more complex interaction model. When asked about why they have considered the documentary videos interaction models harder to use, users argued that it was not due to the designed interaction models/interfaces but due to their cognitive state. In fact, while watching documentary videos, they felt a certain pressure to choose topics of interest even without the need to know more about them, which didn't occur while watching news videos. Thus being, they perception of complexity was associated to the pressure of using the system and to the interaction model/interfaces designed which they have considered easy to use (27–90%) and learn (26–87%).

As a whole, the eiTV crossmedia system was evaluated and the results are presented in Table 6. As may be seen, G1 and G2 results are similar which is a very good indicator. However, and as expected, was better in G1, the group that already participated in

Table 5. Results for the three different video types: CSI, Documentaries and News

The possibility to choose topics	CSI series	Documentaries	News
Where was more useful?	6 (20%)	**15 (50%)**	9 (30%)
Where you used it more?	7 (23%)	10 (33%)	**13 (43%)**
Where you use it less?	**15 (50%)**	10 (33%)	5 (17%)
Where was easier to use?	**19 (63%)**	8 (27%)	3 (10%)
Where was harder to use?	6 (20%)	**13 (43%)**	11 (37%)

previous evaluation and thus developed a certain sense of belonging with the system and a closer relationship with the evaluation team. The results achieved from G2, the group that never interacted with the application, were good and optimistic. In terms of qualitative evaluation, when asked why not interested in having the system: 2 users answered that they usually don't feel the need to know more about topics, 1 user said that he don't like to interact while watching videos and 1 user said that he may use it from time to time.

Table 6. Overall Evaluation of the Whole eiTV System

Whole Application	G1	G2
Useful	14 (93%)	13 (87%)
Easy to use	13 (87%)	11 (73%)
Easy to learn	14 (93%)	12 (80%)
Like to have it	14 (93%)	11 (73%)
Recommend to a friend	15 (100%)	11 (73%)

Important to refer that educational tests were not carried out. As mentioned previously, the generated web-content aggregate additional information based on the users' topics of interest so, they may be considered informal learning environments. Thus being, assessing how learning was improved by this portal was outside the scope of our work.

6 Conclusions and Future Work

Our posed research question was: *is it possible to create an easy, intuitive, natural, and non-intrusive interaction model that allows interacting with news videos in the context of the previous developed eiTV system in an integrated manner?* From the evaluation results obtained, we can conclude that the answer is yes. Both, the interaction models proposed (a1 and a2), to create the personalized web-content from news videos, were considered easy, intuitive, and non-intrusive of the viewing experience. On the other hand, the designed interaction model also allowed interacting with news videos in the context

of the previously developed eiTV and in an integrated manner. In this work, only the create functionality was implemented, for videos about news, in high fidelity prototypes, through the three devices (iTV. Mobile and PC), with the three information levels (1, 2 and 3) available onscreen. This interaction model interfaces were then integrated with previously developed work, namely, eiTV systems with videos about documentaries and CSI series. The final eiTV system prototype, with three types of videos, was then tested. The evaluation results achieved were considered very optimistic. In many aspects, the designed functionalities and the system flexibility were perceived as useful, easy to use and learn, not intrusive or disturbing of the viewing experience and an added value in the crossmedia research area. The designed options allowed to accommodate the changes in users' cognitive mode (e.g., information levels), and the prototypes were designed and tested in a real scenario and context of use with all the devices involved. Finally, considering the design framework followed, the trends and convergence in the use of multiple devices, and the results of this and previous studies, we have reasons to believe that our goal for this crossmedia system is worth pursuing and that we can achieve quite good results in different contexts and scenarios. However, this study presents some limitations: considering that the context of use was limited to a living room, the mobility of the system was not tested; only the create functionality was implemented and tested and, as to the evaluation results, not all the differences between groups were statistically significant, meaning that some results should be considered as an indicator. As to future work, we intend to improve some details about the interfaces and interaction models, implement and test the system as a whole unit (with all the functionalities, running from all the devices and for the three types of videos: news, documentaries, and CSI series) in different contexts. We also intend to continue exploring the devices technological advances to create new input solutions and functionalities capable to better support users' needs, and different cognitive modes, in different scenarios and contexts. Exploration of the system's potential for formal educational settings is also in our plans. The continuous improvement of the interfaces and navigation model, so they may become easier to learn and be used by an elderly population, is also a goal. Finally, we intend to evaluate our interaction model considering the recent findings about operation latency in the context of crossmedia systems [11] already mentioned in the related work section.

References

1. Moloney, K.: Multimedia, Crossmedia, Transmedia… What's in a name? Published on 21 of April (2014). https://transmediajournalism.org/2014/04/21/multimedia-crossmedia-transmedia-whats-in-a-name/. Accessed 23 Sept 2023
2. Gambarato, R.: Crossmedia, Multimedia and Transmedia (2020) Published on 20 of October 2020. https://www.youtube.com/watch?v=G3wdbajO6js. Accessed 23 Sept 2023
3. Jenkins, H.: Transmedia missionaries: Henry Jenkins. Published in 23 of July 2009 (2009). http://www.youtube.com/watch?v=bhGBfuyN5gg. Accessed 20 Sept 2023
4. Prata, A., Chambel, T., de Abreu, J.F.: Personalized web-based contents, generated through a cross-media environment, as additional information to documentary videos. In: Abásolo, M.J., de Castro Lozano, C., Olmedo Cifuentes, G.F. (eds.) Applications and Usability of Interactive TV. jAUTI 2022. Communications in Computer and Information Science, vol. 1820, pp. 3–19. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-45611-4_1

5. Prata, A., Chambel, T.: Personalized interactive video-based crosmedia informal learning environments from iTV, PC and mobile devices – the design challenges. In: Abásolo, M.J., Abreu, J., Almeida, P., Silva, T. (eds.) Applications and Usability of Interactive TV. jAUTI 2020. Communications in Computer and Information Science, vol. 1433, pp. 89–104. Springer, Cham. (2021). https://doi.org/10.1007/978-3-030-81996-5_7
6. Bonometti, S.: Learning in cross-media environment. Int. J. Web-Based Learn. Teach. Technol. **12**(4), 48–57 (2017). https://doi.org/10.4018/IJWLTT.2017100105
7. Brudy et al.: Cross-device taxonomy: survey, opportunities and challenges of interactions spanning across multiple devices. In: CHI 2019: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–28 (2019). https://doi.org/10.1145/3290605.3300792
8. Cardoso, B.: A Unificação no Consumo de Conteúdos Audiovisuais: contributos para a experiência de utilização e sugestões para operadores. In: PhD thesis, Communications and Arts Department, Aveiro university, on 14 January 2022, Portugal (2022)
9. Abreu, J., Almeida, P., Velhinho, A., Varsori. E.: Returning to the TV screen. In: Managing Screen Time in an Online Society, pp. 146–171. IGI Global (2019). https://doi.org/10.4018/978-1-5225-8163-5.ch007
10. Taplin, J.: Long time coming: has interactive TV finally arrived?. opening keynote. In: Proceedings of 9th European Conference on Interactive TV and Video: Ubiquitous TV (EuroiTV'2011), in coop with ACM, p. 9, Lisbon, Portugal (2011)
11. Liu, X., Zhang, Y., Yan, Z., Ge, Y.: Defining 'seamlessly connected': user perceptions of operation latency in cross-device interaction. Int. J. Hum.-Comput. Stud. 177, 103068 (2023). https://www.sciencedirect.com/science/article/abs/pii/S1071581923000770?via%3Dihub
12. Prata, A., Chambel, T., Guimarães, N.: Personalized content access in interactive TV-based cross media environments. In: Kompatsiaris, Y., Merialdo, B., Lian, S., (eds.), TV Content Analysis Techniques and Applications, pp. 331–368. CRC-PRESS, Taylor & Francis Group, March 2012 (2012). ISBN: 978–1–43985–560–7 http://www.crcpress.com/product/isbn/9781439855607;jsessionid=a4luPP5knC7A-jRg5m3f1g**
13. Prata, A., Chambel, T.: The Design of Flexible Video-Based Crossmedia Informal Learning Contexts Beyond iTV, to be published soon in the Elsevier Entertainment Computing Journal (2013)
14. Prata, A., Chambel, T.: Mobility in a crossmedia environment capable of generating personalized informal learning contents from iTV, PC and mobile devices. In: Proceedings of JAUTI 2019 – VIII Conferência Iberoamericana sobre Aplicações e Usabilidade da TV Interativa, pp. 59–71, Rio de Janeiro, Brasil (2019)
15. Prata, A., Chambel, T.: Mobility in crossmedia systems, the design challenges that need to be addressed. In: Abásolo M., Kulesza R., Pina Amargós J. (eds.) Applications and Usability of Interactive TV. jAUTI 2019. Communications in Computer and Information Science, vol. 1202. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-56574-9_5
16. Prata, A., Chambel, T., Abreu, J.: Cross-media environment to generate personalized web-based contents as additional information to documentary videos. In: Proceedings of the 11th Iberoamerican Conference on Applications and Usability for Interactive TV (jAUTI'2022), pp. 15–27. Cordoba, Spain (2022). https://sedici.unlp.edu.ar/handle/10915/162897
17. Prata, A., Chambel, T., Abreu, J.: Personalized extra information to documentary videos, generated through an iTV crossmedia system – improved interfaces. In: Proceedings of the 12th Iberoamerican Conference on Applications and Usability for Interactive TV (jAUTI'2023). Havana, Cuba (2023)
18. Pemberton, L., Fallahkhair, S.: Design issues for dual device learning: interactive television and mobile phone. In: Proceedings of 4th World Conference on mLearning - Mobile Technology: the future of Learn in your hands (mLearn'2005), Cape Town, South Africa (2005)

19. Obrist, M., Knoch, H.: How to Investigate the quality of user experience for ubiquitous TV?. tutorial. In: Proceedings of EuroiTV'2011, 9th European Conference on Interactive TV and Video: Ubiquitous TV, Lisbon, Portugal (2011)
20. Martin, R., Holtzman, H.: Newstream. a multi-device, cross-medium, and socially aware approach to news content. In: Proceedings of the 8th European Interactive TV Conference (EuroiTV 2010), in coop with ACM, pp. 83–90, Tampere, Finland (2010)
21. Abreu, J.: Design de Serviços e Interfaces num Contexto de Televisão Interactiva. In: PhD Thesis, Communications and Arts Department, Aveiro University, Portugal (2007)
22. Livingston, K., Dredze, M., Hammond, K., Birnbaum, L.: Beyond broadcast. In: Proceedings of ACM IUI'2003, The Seventh International Conference on Intelligent User Interfaces, pp. 260–262, Miami, USA (2003)
23. Norman, D.: Things that Make us Smart. Addison Wesley Publishing Company (1993)
24. Segerståhl, K.: Utilization of pervasive IT compromised? understanding the adoption and use of a cross media system. In: Proceedings of 7TH International Conference on Mobile and Ubiqitous Multimedia (MUM'2008) in cooperation with ACM SIGMOBILE, pp. 168–175, Umea, Sweden (2008)
25. Norman, D.: The Design of Everyday Things. Basic Books, New York (2002)
26. Segerståhl, K., Oinas-Kukkonen, H.: Distributed user experience in persuasive technology environments. In: de Kort, Y., IJsselsteijn, W., Midden, C., Eggen, B., Fogg, B.J. (eds.) Persuasive Technology. PERSUASIVE 2007. Lecture Notes in Computer Science, vol. 4744. Springer, Berlin (2007) .Heidelberg. https://doi.org/10.1007/978-3-540-77006-0_10
27. Florins, M., Vanderdonckt, J.: Graceful degradation of user interfaces as a design method for multiplatform systems. In: Proceedings of the ACM International Conference on Intelligent User Interfaces (IUI 2004), pp. 140–147, Funchal, Madeira, Portugal (2004)
28. Nielsen, J.: Coordinating user interfaces for consistency. In: Neuauflage 2002 ed., the Morgan Kaufmann Series in Interactive Technologies, San Francisco, CA, USA (1989)
29. Abreu, J., Almeida, P., Silva, T.: A UX evaluation approach for second-screen applications. In: Communications in Computer and Information Science, vol. 605, pp. 105–120 (2016). https://doi.org/10.1007/978-3-319-38907-3_9
30. Lund, A.: Measuring Usability with the USE Questionnaire (2001). https://garyperlman.com/quest/quest.cgi?form=USE. Accessed 08 June 2023
31. NASA. NASA TLX Paper and Pencil Version (2019). https://humansystems.arc.nasa.gov/groups/tlx/tlxpaperpencil.php. Accessed 23 Sept 2023

Health and Socialization Through Interactive TV

Physical Functionality in Older Adults Through Interactive Digital TV Intervention

María Magdalena Rosado[1,2(✉)], Álvaro Espinoza Burgos[3], María José Abásolo[1], José Antonio Valle Flores[2], Telmo Silva[4], Stalin Jurado[2], and Sheyla Villacres[2]

[1] Schoolof Computer Science, National University of La Plata (UNLP), La Plata, Argentina
maria.rosadoa@info.unlp.edu.ar, mjabasolo@lidi.info.unlp.edu.ar
[2] Facultyof Medical Sciences, Catholic University of Santiago de Guayaquil, Guayaquil, Ecuador
{Stalin.jurado,jose.valle,sheyla.villacres}@cu.ucsg.edu.ec
[3] Faculty of Physical Education, University of Guayaquil, Guayaquil, Ecuador
alvaro.espinozabu@ug.edu.ec
[4] Universityof Aveiro, Aveiro, Portugal
tsilva@ua.pt

Abstract. The geriatric population worldwide is rapidly growing, accompanied by an increase in chronic diseases and disabilities. Rehabilitation plays a crucial role in maintaining and improving independence and quality of life in this group. This study focused on determining the use of interactive digital television as a rehabilitation tool to prevent and/or reduce the risk of falls in older adults.

Functional gait tests were used to identify the type of fall risk. To evaluate the emotions while using interactive television application the Self-Assessment Manikin was used. The body composition was measured through anthropometric parameters.

After 29 exercise sessions, positive impacts on physical function and mood were observed, with significant improvements in flexibility, muscle strength, balance, and gait. These results suggest that interactive digital television could be a useful platform to improve the quality of life of older adults.

Keywords: Plex · Digital Television · Exercises · Geriatric Assessment · Anthropometric Parameters · Older Adults

1 Introduction

The aging process involves various biochemical, physiological, morphological, social, psychological, and functional changes that affect the autonomy, functional independence, and quality of life of older adults [1]. It is associated with the loss of neuromuscular function and performance, partly related to the reduction of muscle strength and power [2, 3].

Therefore, functionality in older adults is crucial for active aging programs [4]. Physical activity is fundamental to maintain a healthy lifestyle and positively affects

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 51–65, 2024.
https://doi.org/10.1007/978-3-031-70439-0_4

overall health and well-being [5]. Evaluating the ability to perform activities of daily living (ADLs) such as climbing stairs, lifting objects, or walking is essential in geriatric assessment [6–8]. Population aging will increase health problems, especially those related to functionality [9]. The demand for physical therapy and rehabilitation services is increasing, especially among older adults [10]. Given the growing aging population, it is crucial to explore new interventions to improve physical function, prevent injuries, and reduce the risk of chronic diseases in older adults [11, 12].

Assistive technologies [1] are an important tool to address the challenges of aging, offering opportunities to improve healthcare and promote healthy lifestyles on a larger scale [5]. Furthermore, these innovative solutions [2] such as virtual reality, artificial intelligence, and telecare can overcome the limitations of access and costs associated with in-person therapy, such as geographical distance or high costs [11].

Alongside the rest of the world, Ecuador is also facing an increase in the number of older adults. This phenomenon presents significant challenges in accessing medical care [12]. In this scenario, home exercise programs can reduce frailty and fall risk, and improve the quality of life of older adults [13–15]. In current study, we have created an interactive training program that allows older adults to continue treatment at home using interactive digital television. With the tool described in this paper, individuals can perform exercises evaluated and constructed by physical therapists, complementing follow-up and evaluation of their progress.

This article presents the research results and is organized as follows: Sect. 2 describes the materials and methods used in the study. Section 3 presents the experimental findings obtained. Section 4 argues the discussion of these findings. Finally, the conclusions of the study are presented in Sect. 5.

2 Methods

2.1 Participants

The elderly who attend the Dr. Arsenio de la Torre Marcillo Gerontological Center, which has an agreement with the Catholic University of Santiago de Guayaquil, were invited to participate in the study. The aim of the study was to prevent and/or reduce the risk of falls using an interactive digital television application.

To select the participants, inclusion and exclusion criteria were defined by two of the researchers, both physiotherapists. Between May and June 2023, elderly individuals from this center were recruited, and evaluations were conducted in June 2023. All participants provided written informed consent before joining the study and were selected according to a set of detailed criteria presented in Table 1.

Initially, twenty-five elderly individuals were recruited, but only sixteen independent ones, aged between 65 and 84 years with possible mobility disorders, met the established selection criteria.

2.2 Measurement Procedures

The entire evaluation and follow-up process was carried out by physical therapy faculty members with more than five years of experience to reduce biases [1]. A medical record was created where personal data, medical, and family history were requested.

Table 1. Criteria for the selection

Inclusion	Exclusion
- Able to walk independently - Follow simple instructions - No severe cognitive disorders -Have a Smart TV with Android operating system - Internet service	-Auditory/visual impairments that prevent interaction with the Smart TV - Hemispatial neglect - Ataxia - Unresolved traumatic injuries such as fractures - Comprehensive memory impairments

Each participant (Fig. 1) underwent anthropometric measurements: weight (kg), height (cm), and body mass index (BMI). A calibrated digital scale and stadiometer were used, following established protocols, to calculate BMI using the formula weight (kg)/height $(cm)^2$ [16].

For the assessment of fall risk (Fig. 2), the following tests were applied:

i) The Tinetti scale has two domains: gait and balance. Its main objective is to detect the risk of falls. Its predictive value is higher than the value in the muscle test [17, 18]. It consists of two parts: the first is balance, which consists of nine tests: 1) sitting balance; 2) rising; 3) attempts to rise; 4) immediate balance after rising; 5) balance while standing; 6) push; 7) eyes closed; 8) 360-degree turn; and 9) sitting down. The second part, gait, consists of 7 tests: 1) initiation of gait; 2) step length and height; 3) step symmetry; 4) step continuity; 5) path; 6) trunk; and 7) walking posture. The maximum score for the gait subscale is 12, and for balance, it is 16. The sum of both scores determines the risk of falls. When the score is high, the risk of falls is lower; conversely, when it is low, the risk of falls increases [19]. The classification of fall risk was: <= 18 points, high risk of falls; between 19–24, moderate risk of falls; and >24, low risk of falls [20].
ii) The timed Get Up and Go test is an auxiliary test in the diagnosis of gait and balance disorders and their association with a determined risk of falls. The application time is 10 min, and the classification of fall risk was: >13 s, high risk of falls; between 11–13, moderate risk of falls; and <10 s, low risk of falls [21].

To evaluate emotions concerning the digital TV application, the Self-Assessment Manikin (SAM) was used, a tool that measures three emotional dimensions: pleasure, arousal, and dominance. This instrument consists of three groups of pictograms with humanoid figures, one for each dimension of emotion. Each group consists of 5 drawings and four spaces between them, allowing the subject to move within a range of 9 points per dimension [22]. Regarding the meaning of the dimensions:

– Pleasure: refers to how a person experiences emotions (smiling – angry).
– Arousal: determines the intensity or level of energy associated with the mood (calm -aroused).
– Dominance: refers to the feeling a person has about controlling their own emotions at a given moment (low or highly dominated).

a) b)

Fig. 1. Evaluation of anthropometric measurements in older adults: a) Dynamometry; b) Height measurement

a) b)

Fig. 2. Assessment of fall risk in older adults: a) Balance or gait test; c) Administering the timed Up and Go test

2.3 Intervention Protocol

A proprioceptive exercise program focused on improving strength, muscular endurance, coordination, and stability for fall prevention in older adults was implemented. The program lasted 29 sessions, divided into five phases:

- **First phase:** Socialization. A meeting was held with all older adults in the Biomechanics laboratory of the Faculty of Health Sciences at the Catholic University of Santiago de Guayaquil, Ecuador. During this meeting, the research and program were introduced.
- **Second phase:** Initial evaluation. Anthropometric measurements were taken, and the Tinetti and Timed Up and Go scales were applied.
- **Third phase:** Adaptation. A forty-minute adaptation session with the tool was conducted.
- **Fourth phase:** Intervention. This phase consisted of twenty-nine work sessions that included flexibility, muscle strengthening, and balance exercises. At the beginning of

each session, a warm-up with joint movements was performed, and at the end, static stretches of the main muscle groups were done. Each session lasted 30 min, distributed among different types of training. The first fifteen sessions were carried out in the university laboratory. During the exercise, the older adults were assisted by a physiotherapist, who provided postural corrections and feedback when necessary. At the end of the first session, the participants' emotional response was evaluated using the Self-Assessment Manikin (SAM). Instructions were given explaining the objective and how to assess the emotions of Pleasure, Excitement, and Dominance, strictly following an established protocol [23]. Subsequently, another fourteen work sessions were conducted at the participants' homes, following the protocol used during the intervention phase.

– **Fifth phase:** Final evaluation. The same data collection instruments as in the second phase were applied, and additionally, the Self-Assessment Manikin test.

Table 2 shows the rehabilitation training macrocycle, which lasted six months. It describes the different intervention phases and the corresponding training volume for each phase.

Table 2. Planning of the intervention program

Month	June				July				August				September				October				November				
# Month	1				2				3				4				5				6				
Week Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Week Start Date					1	8	15	22	29	5	12	19	28	4	11	18	25	2	9	16	23	30	6	13	Total
Week End Date					7	14	21	28	4	11	18	25	3	10	17	24	1	8	15	22	29	5	12	19	
Days W x Week						2	2	2	2	2	2	2	1			2	2	2	2	2	2	2			
Sessions x Week						2	2	2	2	2	2	2	1			2	2	2	2	2	2	2			29
Hours x Week						1	1	1	1	1	1	1	0,5			1	1	1	1	1	1	1			14,5
(1)Socialization Phase	✓																								
(2)Initial Assessment Phase	**Pre test**																								
(3)Adaptation Phase					✓																				
(4)Intervention Phase						✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓			
(5)Final Assessment Phase																									**Post test**
																									Tot. Vol. (meso(min)
Meso Volume (min)									480											420					900
Weekly Volume (min)						60	60	60	60	60	60	60	60			60	60	60	60	60	60	60			900
(1)Adaptation Phase					100%																				
(4)Intervention Phase									100% (Institution)								100%(Spatial)								

W: weekly; Vol.: volume; hr: hour; meso: mesocycle

Source: Adaptation of the planning model of Matvéiev [24]

2.4 Intervention and Control

The participants interact with the PLEX application interface to perform the training tasks with the 9 exercise videos on smart TV screens. Priority was given to the fluidity of navigation and selection of exercise videos on the TV, as well as the ergonomic design of the Chromecast remote control. For the first fifteen work sessions with the PLEX application, we installed a 32-inch Riviera TV, model RLED-DSG32HIK5600, equipped with an Android operating system, in the biomechanics laboratory of our

institution. In collaboration with the Faculty's IT support area, configuration tests were carried out (Fig. 3).

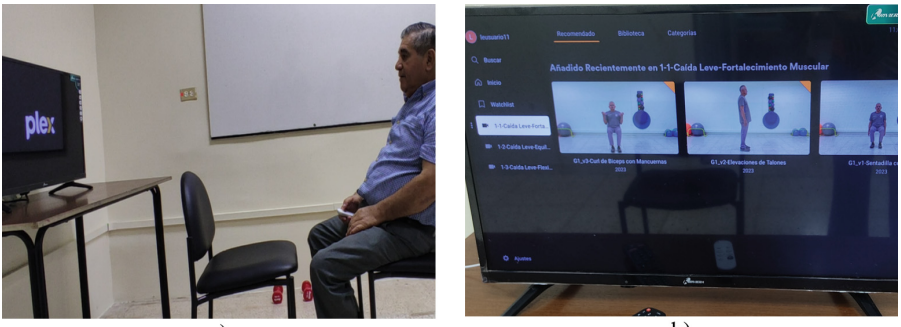
a) b)

Fig. 3. Implementation of the tool for smart TVs: a) Plex installed on the smart TV; b) User access

These tests were performed in various cases and scenarios to ensure proper access to the library of adapted and personalized exercises for each older adult, according to the diagnosis of their functional evaluation, as well as to verify the accessibility of the videos hosted on a Network Attached Storage (NAS).

Older adults received individual instructions on how to use the tool during a forty-minute adaptation session, days before starting the first fifteen rehabilitation treatment sessions (Fig. 4). Subsequently, to continue the program, Plex was installed in each individual's home in the fifteenth week, allowing them to continue with rehabilitation during the remaining fourteen sessions.

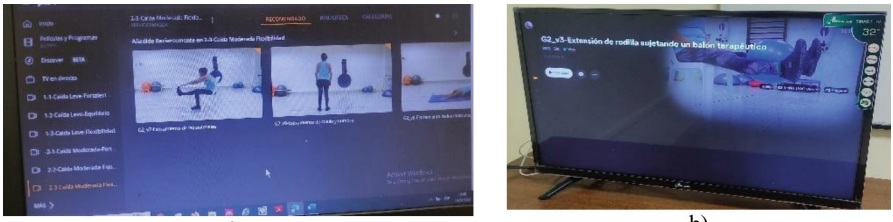
The intervention did not have direct supervision. However, telephone follow-up was carried out with the participants at the end of weeks 17 to 22 to facilitate compliance and resolve any difficulties with the tool. If difficulties were encountered with accessing the exercise video library in Plex, the possibility of additional home visits was offered. Nevertheless, this option was never required due to all participants' willingness to use the tool.



a) b)

Fig. 4. Personalized multimedia platform for smart TVs: a) Patient using Plex on smart TV; b) Access to muscle-strengthening exercises on a smart TV

The intervention consisted of 27 exercise videos, distributed into 9 categories according to the risk of falls (mild, moderate, and severe), in MP4 format with a resolution of 1920 × 1080 pixels and using the MPEG-4 video codec. To optimize understanding and learning, instructional scripts were created for the videos, prioritizing the clarity and effectiveness of the instructions (Fig. 5).



a) b)

Fig. 5. Exercise platform to prevent falls: a) Plex exercise library on smart TV; b) Smart TV showing therapy ball knee extension exercise on Plex

2.5 Exercise Design

Based on a review of scientific articles [25, 26], exercises that were effective in reducing and preventing the risk of mild, moderate, and severe falls to improve an individual's functional capacity were established. These exercises are designed to increase strength, balance, coordination, and stability in older adults, following a 30-min exercise protocol during 29 sessions, twice a week.

Exercises for mild falls are designed for situations where there is less risk of falling, such as tripping over a small object or slightly uneven ground. These may include simple actions that improve posture and balance awareness, such as standing on one leg, walking in a straight line with eyes closed, heel raises, among others, as illustrated in (Fig. 6).

Exercises for severe falls are designed for individuals who have already experienced a fall or have a high risk of falling. These exercises should be performed under the supervision of a physical therapist or healthcare professional. These may include resistance and strength training to strengthen the core and leg muscles, to reduce the risk of injuries, as shown in (Fig. 8).

Exercises for moderate falls are designed for situations where there is a greater risk of falling, such as losing balance while walking on a wet or uneven surface. Other exercises are illustrated in (Fig. 7).

3 Results

The statistical package SPSS V22 was used to obtain descriptive statistics, measures of central tendency (mean, minimum, maximum, and mode), and measures of dispersion (standard deviation, minimum, and maximum). Two non-parametric statistical tests were applied (Wilcoxon and Friedman), which are used to work with small population samples [27].

The Wilcoxon test was used to compare two related variables when the data was not normally distributed. This test was used to compare the group before and after the intervention (i.e., related samples) with the Tinetti scale and the Get Up and Go [28].

The Friedman test was used to compare three or more related variables (pleasantness, emotion, and control). This test examined the ranks of the data generated in each time period that was evaluated with the SAM test to determine if the variables share the same continuous distribution of their origin [29].

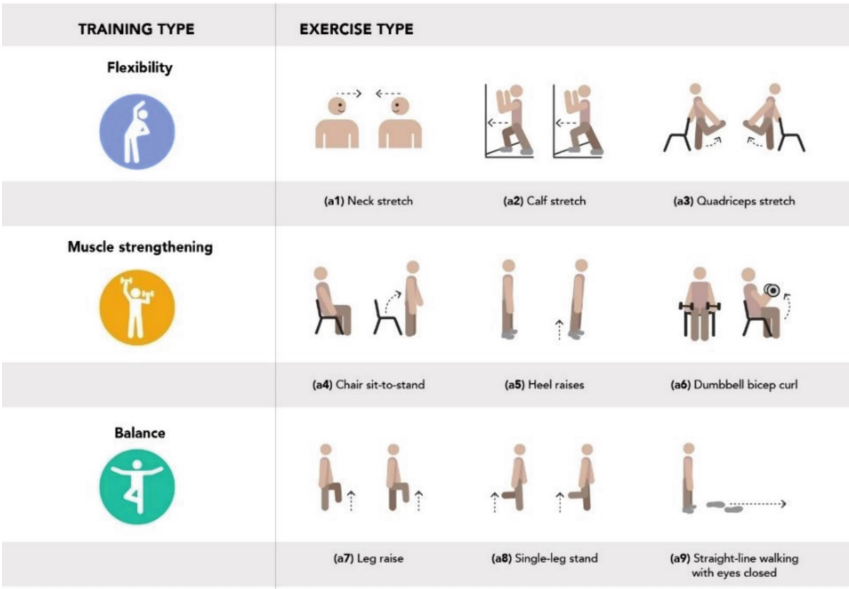


Fig. 6. Description of exercises for the patient with a mild risk of falls

A significance level of $p < 0.05$ was considered for both statistical tests. The study population initially consisted of sixteen older adults of both sexes who participated in the 15 baseline sessions conducted in the Biomechanics laboratory. However, only ten older adults continued with the exercise program in the 14 subsequent sessions scheduled at home. Therefore, the results obtained are only based on the group that completed all the routines from the first to the last session of the program.

Table 3 shows the body composition before and after using digital television for physical exercise. A reduction was observed in the values of: body weight (68.65 ± 11.87 vs 65.13 ± 12.10; $\Delta = -3.49$; $p > 0.013$), waist circumference (89.28 ± 7.22 vs 84.40 ± 7.67; $\Delta = -4.88$; $p < 0.003$), BMI (29.23 ± 3.56 vs 26.43 ± 4.31; $\Delta = -2.80$; $p > 0.007$), body fat (38.78 ± 6.71 vs 34.69 ± 7.90; $\Delta = -4.09$; $p > 0.007$) and visceral fat (1.63 ± 0.90 vs 1.63 ± 0.68; $\Delta = 0.00$; $p < 0.005$). However, an increase in lean mass was observed (64.16 ± 11.82 vs 67.30 ± 13.77; $\Delta = +3.14$; $p > 0.011$), as detailed in Table 3.

Regarding the results of Table 4 on the assessment of balance and gait using the Tinetti Scale, it was observed that the mean was 22.40 in the initial assessment and 26.20 in the

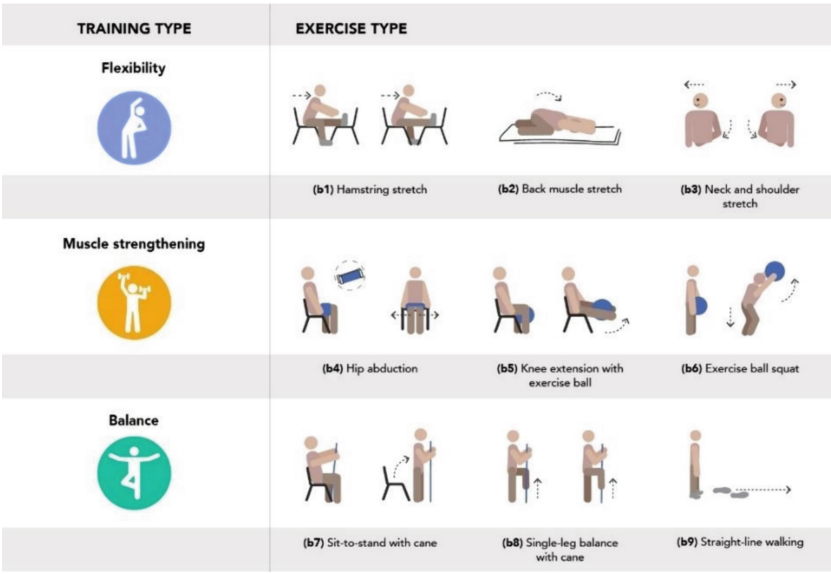


Fig. 7. Description of exercises for the patient with a moderate risk of falls

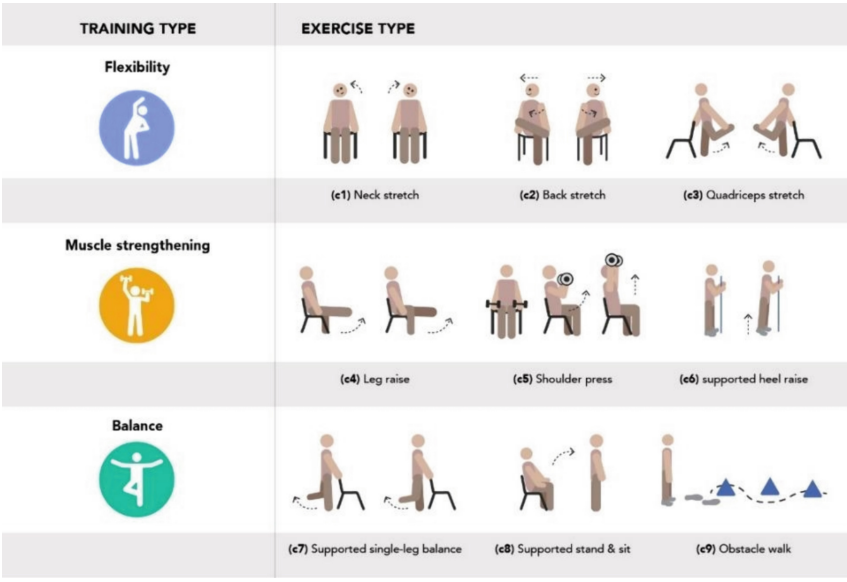


Fig. 8. Description of exercises for patients at risk of severe falls

final assessment. The standard deviation was 4.14 in the initial assessment and 2.49 in the final assessment. The minimum score recorded was 18 in the initial assessment and 22 in the final assessment, while the maximum was 28 in both assessments. In relation

to the results of the Get Up and Go test, it was observed that the mean time was 12.09 s in the initial assessment and 9.92 s in the final assessment. The standard deviation was 1.94 s in the initial assessment and 1.64 s in the final assessment. The minimum time recorded was 9.76 s in the initial assessment and 7.22 s in the final assessment, while the maximum time was 15 s in the initial assessment and 13 s in the final assessment.

Table 3. Body composition before and after using interactive digital television.

Variable	Before		After		
	Mean ± SD	Min - Max	Mean ± SD	Min - Max	P-value*
Weight (Kg)	68,65 ± 11,87	42,15–82,35	65,16 ± 12,10	40,7–78,40	0,013
Waist Circumference (cm)	89,28 ± 7,22	78–100	84,40 ± 7,67	70,2–93,30	0,003
BMI (Kg/m^2)	29,23 ± 3,56	22,80–34,28	26,43 ± 4,31	20,20–33,30	0,007
Body Fat (%)	38,78 ± 6,71	28,50–49,70	34,69 ± 7,90	18,20–45,40	0,007
Visceral Fat (%)	1,63 ± 0,90	1,20–3,70	1,63 ± 0,68	0,77–2,50	0,005
Lean Mass (%)	64,16 ± 11,82	50,30- 91,90	67,30 ± 13,77	53,50–97,60	0,011

SD: standard deviation

Table 5 presents the classification of the variable "fall" at two different time points. In the initial evaluation, it will be ensured that 40% of older adults are classified as having a severe risk of falls, 30% with a moderate risk of falls, and 30% with a mild risk of falling. However, in the final evaluation, a significant change was observed: 70% of the participants shifted to a mild risk of falls, while the remaining 30% remained at a moderate risk of falls. It is important to note that no older adults without a severe risk of falls were recorded in either evaluation. These results suggest that the interventions or treatments implemented between the initial and final evaluations have had a positive impact on reducing the risk of falls in the studied population of older adults. The hypothesis contrast test between the initial and final fall risk, performed with the Wilcoxon non-parametric test for two related samples, showed a statistical significance of $p < 0.05$. Therefore, the null hypothesis is rejected, which indicates that the use of the exercise program using the application with interactive digital television in static and dynamic balance is effective in preventing falls in older adults who participated in the study.

Table 6 shows the classification of the variable "fall" in the initial assessment. Before the intervention, it is observed that 40% of older adults were classified in the category of severe risk of falls, while 20% were at moderate risk and another 40% at mild risk of falls. This indicates an equitable distribution among the three risk levels before the intervention. After the intervention, a notable change in the distribution of older adults across different risk levels is observed. 80% of the participants shifted from being at mild risk of falls before the intervention to remaining in that category afterward. Additionally, the 20% who were at moderate risk of falls before the intervention remained in that

category afterward. However, it is important to note that no older adults were registered in the severe risk category of falls after the intervention. The hypothesis contrast test between the initial and final fall risk, performed with the Wilcoxon non-parametric test for two related samples, showed a statistical significance of $p < 0.05$. Therefore, the null hypothesis is rejected, which indicates that the use of the exercise program using the application with interactive digital television in mobility and functionality is effective in preventing falls in older adults who participated in the study.

Finally, to assess the level of satisfaction with the use of digital television through the emotional responses of the 10 older adults, statistically significant differences were observed in three stimuli: pleasure, dominance, and arousal ($p < 0.05$), using the PLEX application. Further details are provided in Table 7.

Table 4. Results of the Initial and Final Assessment of Fall Risk, according to the Tinetti Scale and Get Up and Go with the use of interactive digital television

Descriptives	Initial Tinetti Scale	Final Tinetti Scale	Initial Get Up And Go Test	Final Get Up And Go Test
Mean	22,40	26,20	12,09	9,92
Mode	18	28	15	10
SD	4,14	2,49	1,94	1,64
Minimum	18	22	9,76	7,22
Maximum	28	28	15	13

Table 5. Results of the Initial and Final Assessment with the Tinetti Scale

Score	Diagnosis	Before	After	P-value*
>24 points	Low Fall Risk	3[30]	7[70]	0,007
19–24 points	Moderate Fall Risk	3[30]	3[30]	
<+ 18 points	Severe Fall Risk	4[40]	0[0]	

Table 6. Results of the Initial and Final Assessment with the Get Up and Go Test

Score	Diagnosis	Before	After	P-value*
<10 sg	Low Fall Risk	4[40]	8[80]	0,005
11 a 13 sg	Moderate Fall Risk	2[20]	2[20]	
>13 sg	Severe Fall Risk	4[40]	0[0]	

Table 7. Results of Emotional Responses Before and After Using the Self-Assessment Manikin Test

Emotional Responses	Before		After		
	Mean ± SD	Min - Max	Mean ± SD	Min - Max	P-value*
Pleasantness Stimulus	7,20 ± 1,75	5,00–9,00	8,60 ± 0,516	8,00–9,00	0,026
Control Stimulus	7,88 ± 1,23	5,00–9,00	9,00 ± 0,00	8,00–9,00	0,016
Activation Stimulus	7,00 ± 1,33	5,00–9,00	8,70 ± 0,48	8,00–9,00	0,011

4 Discussion

The purpose of this study was to investigate how interactive digital television can be used by older adults to perform exercises to prevent or follow up on rehabilitation treatment to reduce the risk of falls.

The study focused not only on evaluating how participants manage this technology but also on measuring changes in their anthropometric measurements before and after participating in the intervention exercise programs.

A reduction was observed in body weight (3.49 kg), waist circumference (4.88 cm), BMI (2.80 kg/m^2), body fat (4.09%), and visceral fat (0.00%). However, an increase in musculoskeletal mass (3.14%) was observed.

Although studies on different technological tools and changes in body composition are limited, research conducted in older adults where the effect of physical activity programs on body composition is measured has indicated that long-term muscle strength training can be effective in improving muscle mass, functional strength, physical capacity, and reducing fat mass in older adults [30].

On the other hand, according to the study by [31], exergames, also known as digital exercise games, improve various aspects of gait in older adults. During a 12-week training program, participants engaged in interactive activities that challenged their balance and coordination. Variables related to gait (speed, cadence, time, step length, and step width) were assessed before and after training. The results showed significant improvements in walking after the training program. Additionally, positive changes in body composition were observed, with a reduction in body fat and an increase in fat-free mass after training.

Using the Self-Assessment Manikin (SAM) test in our study, we observed an increase in the emotion experienced in the dimensions of "pleasantness," "control," and "activation" before and after using the PLEX application (7.20 vs 8.60; 7.88 vs 9.00; 7.00 vs 8.70 respectively).

The results obtained in this study are similar to previous research carried out in the same population, although using different technological tools. For example, in the study "Games for a Better Life: Effects of Playing Wii Games on the Well-Being of the Elderly in a Long-Term Care Facility", the SAM was used to evaluate the health effects of older adults participating in activities with the Wii console in nursing homes. After 8 weeks of participation, an improvement in positive emotions and a decrease in negative emotions was observed [32].

In another pilot intervention study titled "Exercise Games for Subsyndromal Depression in Older Adults: A Pilot Study of a Novel Intervention" [33], which explored the use of exercise games in older adults with subclinical depressive symptoms, the SAM was also used to assess emotional state. It was found that after performing exercises with interactive games of dance, baseball, bowling, tennis, and boxing on the Xavix console, there was a significant improvement in the evaluation of pleasure and arousal according to the SAM.

5 Conclusion

This study focused on determining the use of interactive digital television as a rehabilitation tool to prevent and/or reduce the risk of falls in older adults. There were multiple significant benefits for older adults who participated in the exercise program designed for fall risk prevention, using the PLEX application interface. Firstly, a notable reduction was observed in body weight, waist circumference, body mass index (BMI), body fat, and visceral fat, while there was an increase in lean muscle mass. Additionally, the evaluations of the Tinetti Scale and the Get Up and Go test evidenced a substantial improvement in balance and gait in the participants. A key finding was the significant decrease in the risk of falls, with 70% of older adults shifting to a mild risk after the intervention.

Statistically significant differences were observed in the levels of pleasure, dominance, and emotional arousal of the participants, as measured by the Self-Assessment Manikin, suggesting an improvement in their emotional state and satisfaction with the use of interactive digital television for performing physical exercises focused on fall prevention.

However, it is crucial to acknowledge some limitations of the present study. The sample was limited to a few participants, which restricted the generalization of the results. Additionally, it focused on a population with access to technology, which may not be representative of all elderly communities. Future research should expand the sample size, consider greater demographic diversity, and conduct comparative studies that evaluate different technological and traditional interventions, as well as their long-term effects on fall prevention and quality of life improvement.

References

1. Fuzhong, L.: Physical activity and health in the presence of China's economic growth: Meeting the public health challenges of the aging population. J. Sport Health Sci. **5**(3), 258–269 (2016)
2. Carville, S.F., Perry, M.C., Rutherford, O.M., Smith, I.C., Newham, D.J.: Steadiness of quadriceps contractions in young and older adults with and without a history of falling. Eur. J. App. Physiol. **100**(5), 527–533 (2007)
3. Macaluso, A., De Vito, G.: Muscle strength, power and adaptations to resistance training in older people. Eur. J. App. Physiol. **91**(4), 450–472 (2004)
4. Laguado J., Camargo Hernández, K., Campo Torregroza, E., Martín Carbonell, M.: Funcionalidad y grado de dependencia en los adultos mayores institucionalizados en centros de bienestar. Gerokomos. **28**(3), 135–141 (2017)

5. Kamali, M.E., Angelini, L., Caon, M., Carrino, F., Rocke, C., Guye, S., et al.: Entrenadores virtuales para el bienestar de los adultos mayores: una revisión sistemática. en IEEE Access. **8**, 101884–101902 (2020). https://doi.org/10.1109/ACCESS.2020.2996404
6. Zavala-González, M.A., Domínguez-Sosa, G.: Funcionalidad para la vida diaria en adultos mayores. Rev. Médica del Instituto Mexicano del Seguro Soc. **49**(6), 585–590 (2011)
7. Paterson, D., Warburton, D.: Physical activity and functional limitations in older adults: a systematic review related to Canada's physical activity guidelines. Int. J. Behav. Nutr. Phys. Act. **7**(1), 38 (2010)
8. Segovia Díaz de León, M.G., Torres Hernández, E.A.: Funcionalidad del adulto mayor y el cuidado enfermero. Gerokomos **22**(4), 162–166 (2011). https://doi.org/10.4321/S1134-928X2011000400003
9. Echeverría, A., Astorga, C., Fernández, C., Salgado, M., Villalobos Dintrans, P.: Funcionalidad y personas mayores: ¿dónde estamos y hacia dónde ir? Rev. Panamericana de Salud Pública. **46**(1) (2022). https://doi.org/10.26633/RPSP.2022.34
10. Yáñez-Yáñez, R., Loncon, R., Elizama, V., Mc Ardle-Draguicevic, N., Cigarroa, I.: Efectos de un programa de telerehabilitación sobre la funcionalidad en personas mayores. Horizonte Sanitario **21**(2), 282–290 (2022)
11. Miciak, M.M., et al.: Las condiciones necesarias de compromiso para la relación terapéutica en fisioterapia: un estudio de descripción interpretativa. Arch. Fisioterapia **8**(1), 1–12 (2018)
12. Khan, M.M.R., Sunny, M.S.H., Ahmed, T., Shahria, M.T., Modi, P.P., Zarif, MII., et al.: Desarrollo de un sistema de telerehabilitación asistido por robot con IIoT integrado y gemelo digital. en IEEE Access. **11**, 70174–70189 (2023). https://doi.org/10.1109/ACCESS.2023.3291803
13. Geraedts, H.A., Zijlstra, W., Zhang, W., Bulstra, S., Stevens, M.: Adherence to and effectiveness of an individually tailored home-based exercise program for frail older adults, driven by mobility monitoring: design of a prospective cohort study. BMC Public Health **14**(1), 570 (2014). https://doi.org/10.1186/1471-2458-14-570
14. Stones, G.: At home it's just so much easier to be yourself': older adults' perceptions of ageing in place. Ageing Soc. **36**(3), 449–481 (2016). https://doi.org/10.1017/S0144686X14001214
15. Serino, S., Barello, S., Miraglia, F., Triberti, S., Repetto, C.: Virtual reality as a potential tool to face frailty challenges. Front. Psychol. **8** (2017). https://doi.org/10.3389/fpsyg.2017.01541
16. Molano-Tobar, N.J., Chalapud Narváez, L.M., Villaquirán Hurtado, A.F.: Estimación de obesidad desde índices de adiposidad en universitarios de Popayán, Colombia. Cultura, Ciencia y Deporte. **16**(48), 301–309 (2021). https://doi.org/10.12800/CCD.V16I48.1753
17. Ser Quijano, T., Del Peña-Casanova, J.: Evaluación Neuropsicológica y funcional de la demencia. Editores JR Prous (1994)
18. Tinetti, M.E.: Preventing falls in elderly persons. J Am Geriatr Soc. **34**, 116–119 (1986)
19. Gutiérrez Pérez, E., Meneses Foyo, A., Bermúdez, P., Gutiérrez Díaz, A., Padilla Moreira, A.: Utilidad de las escalas de Downton y de Tinetti en la clasificación del riesgo de caída de adultos mayores en la atención primaria de salud. Acta Médica Del Centro. **16**(1), 127–140 (2022)
20. Carballo-Rodríguez, A., Gómez-Salgado, J., Casado Verdejo, I., Ordás, B., Fernández, D.: Estudio de prevalencia y perfil de caídas en ancianos institucionalizados. Gerokomos **29**(3), 110–116 (2018)
21. Podsiadlo, D., Richardson, S.: The timed Up & Go: a test of basic functional mobility for frail elderly persons. J. Am. Geriatr. Soc. **39**(2), 142–148 (1991)
22. Romo-González, T., González-Ochoa, R., Gantiva, C., Campos-Uscanga, Y.: Valores normativos del sistema internacional de imágenes afectivas en población mexicana: diferencias entre Estados Unidos. Colombia y México. Univ Psychol. **17**(2), 1–9 (2018)

23. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. Gainesville: University of Florida (2008)
24. Matveiev, L.P.: Fundamentos del entrenamiento deportivo. Editorial Paidotribo, Barcelona (1983)
25. Moreno, R., Ramirez, L.A., Párraga, J.A: Programas de ejercicio físico para la prevención de caídas en personas mayores: revisión sistemática, **1**(2), 45–53 (2019)
26. Ortiz Reyes, G. M., Pérez Pérez, D. F., Muyulema Moyolema, D. D. R., & Córdova Velasco, L. E.: Ejercicios de equilibrio y coordinación en el adulto mayor con riesgo de caída. Mediciencias UTA. **5**(4.1), 75–81 (2021). https://doi.org/10.31243/mdc.uta.v5i4.1.1133.2021
27. Gómez-Gómez, M., Danglot-Banck, C., Vega-Franco, L.: Sinopsis de pruebas estadísticas no paramétricas. Cuándo usarlas. **70**(2), 91–9 (2003)
28. Flores-Ruiz, E., Miranda-Novales, M.G., Villasís-Keever, M.Á.: El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. Estadística inferencial. Rev. Alergia México **64**(3), pp. 364–70 (2017)
29. Berlanga Silvente, V., Rubio Hurtado, M.J.: Clasificación de pruebas no paramétricas. Cómo aplicarlas en SPSS. Rev. d'Innovació i Recerca en Educació **5**(2), 101–103 (2011)
30. Fuentes-Barría, H., Urbano-Cerda, S., Aguilera-Eguía, R., González-Wong, C.: Ejercicio físico y suplementación nutricional para el combate de la obesidad sarcopénica en adultos mayores. Universidad de Salud. **23**(1), 46–54 (2021). https://doi.org/10.22267/rus.212301.213
31. Van Roekel, E., Verhagen, E., Engbers, L., Vriens, B., Lok, M., Erler, N.S., Wouters, E.J.: Los efectos de los juegos de ejercicio de plataformas digitales sobre los parámetros espacio temporales de la marcha entre adultos mayores. Informes científicos **11**(1), 1–12 (2021)
32. Jung, Y., Li, K.J., Janissa, N.S., Gladys, W.L.C., Lee, K.M.: Juegos para una vida mejor: efectos de jugar juegos de Wii en el bienestar de las personas mayores en un centro de atención a largo plazo. Actas de la Sexta Conferencia de Australasia sobre entretenimiento interactivo, 1–6 (2009). https://doi.org/10.1145/1746050.1746055
33. Rosenberg, D., et al.: Juegos de ejercicios para la depresión subsindrómica en adultos mayores: un estudio piloto de una intervención novedosa. Rev. estadounidense de psiquiatría geriátrica **18**(3), 221–226 (2010). https://doi.org/10.1097/JGP.0b013e3181c534b5

Alexa, Send a Hug: TV and Virtual Assistants to Empower Older Adults and Stimulate Intergenerational Connections

Juliana Camargo(✉), Telmo Silva, and Jorge Abreu

DigiMedia, Department of Communication and Arts, University of Aveiro, Aveiro, Portugal
{julianacamargo,tsilva,jfa}@ua.pt

Abstract. As is widely known, the world's population is aging, emphasizing the need to integrate older adults into society through the development of specific digital solutions. Within this context, researchers should explore technologies that can help alleviate feelings of loneliness and address social isolation indicators. In addition, if these demographic lacks familiarity with recent technology resources, the exclusion will not stop escalating significantly. This is the scenario that guides this work, which consists of describing the process of creating a prototype aimed at sending (in a proactive way) personalized notifications to the TV-set, that can be easily answered via a combined voice assistant. These notifications have the potential to encourage older adults to contact their families and/or friends, mitigating the usual complexity of these interactions by doing it using only their voice. This paper describes all the process stages, from the initial interviews and technologies involved, to the preparation for the field tests. Additionally, it provides insights from six tests carried out with elderly people aged between 70 and 89 in a real-life context. This experiment demonstrated the potential of proactive notifications to encourage older adults to interact with technology and, consequently, to contact their relatives more often. The tests also showed that voice commands are important resources for eliminating barriers that make intergenerational connections impossible.

Keywords: older adults · elderly · television · TV · iTV · social connections · notifications · voice assistant · Alexa · voice commands

1 Introduction

Although it is a natural process, getting older is a path permeated by difficulties, prejudice and exclusion. Society tends to isolate older citizens [1, 2] – and, unfortunately, this is not a recent problem. In 1970, for example, the French philosopher and activist Simone de Beauvoir [3] already addressed the issue critically and incisively. According to her, instead of society recognizing the maturity and wisdom of these individuals, "we rob them of their confidence, of the possibilities of a path and of meaning". These are words that seem to be extremely relevant today.

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 66–84, 2024.
https://doi.org/10.1007/978-3-031-70439-0_5

In the same literary work, Simone de Beauvoir provokes reflection on a possible way of easing distancing and exclusion. From the author's point of view, an intelligent - and viable - way of reducing ageism and isolation would be to "associate the strength of the young with the experience of the old". In other words, to connect the generations, exchanging knowledge, feelings and/or support in general.

Transported to the 21st century, the author's thinking comes up against something that can be a chasm between the different generations: technology. Not having access to technological devices and digital platforms, especially those that enable social interactions, can contribute to opening up a "gap" between individuals belonging to different age groups [4, 5]. But how can this distancing be reversed? The answer may seem contradictory at first glance, but it is intrinsically linked to technology. Thinking of solutions aimed at older adults, with features that make them easier to use, can encourage them to experiment [6]. Another important stimulus is the support of young people who have the patience and didactics to guide, answer questions and support older adults on an ongoing basis [7]. Such a combination is crucial for empowering older people when it comes to using technology. Consequently, it has proven to be a powerful mechanism for reducing the gaps between generations.

This study takes place precisely in this context - and aims to gather scientific evidence on the role of technology in boosting intergenerational connection. Based on the perception of the target audience (older adults), a system was developed that combines different technological resources. Its purpose is to facilitate exchanges between the generations, whether by text/audio messages, phone calls or physical encounters. The prototype is based on the articulation between a platform that sends notifications to the television and a virtual assistant (its operation is described in detail below).

In order to present the solution - and the preliminary results of the field tests - this paper has been divided into different sections. The first part, in Sect. 2, presents a compilation of related work, which shows the potential of television and voice assistants to promote social interactions between senior citizens and their loved ones. This is followed by a presentation in Sect. 3 of the methodology chosen for the study and details of all the steps required for data collection. Section 4 contains a description of the system and the technologies used. Finally, the results of the initial field trial with six elderly people are presented in Sect. 5, being the conclusions drawn from this experience and expectations for future work presented in Sect. 6.

2 Related Work

The TV is a device that connects people, whether it's to start a conversation or to bring the family together in the living room. The gadget has a significant capacity to stimulate communications between individuals, making a significant contribution to bringing people closer together [8]. An example of this is TV series, which have an important social component: people mobilize to talk about the episodes the next day [9]. However, there are several indications that the device may represent more than that: it has the capacity to reduce the barriers between older adults and digital inclusion. In the study [10], TV-set was used to make it easier for older adults to access health-related information and recent photographs of their relatives. These resources were organized into television channels

that could be accessed via remote control. In total, 14 participants used the system in a real-life context and highlighted that: i) the "service as a channel" paradigm simplified access for the elderly to information available on the internet; ii) all the services tested were appreciated by the group; iii) the channels that promoted some kind of connection with the family were the most highly rated; iv) individuals with well-established habits, such as using a paper diary, were not receptive to using the TV as a diary.

The device's potential to connect generations was also assessed in the study of Tapia, Gutierrez and Ochoa [11]. An application used on the Smart TV (via Google Chromecast) allowed older adults to exchange emails and share photos with their families. Eight participants tested the system in the laboratory and were able to carry out the proposed tasks without ever having interacted with a Smart TV before. After the tests there was a focus group session and, in consensus, the participants stated that they were "enthusiastic about the solution" and that the prototype is a "way to facilitate interaction and access to content produced by their families".

A similar proposal was explored in the study conducted by Coelho, Rito and Duarte [12]. The prototype consists of an adapted version of the Facebook application for television. According to the elderly people who tested the tool, the fact that it was available on TV made it easier for them to use the social network. By being able to access and view content more easily, they said they felt closer to their families.

Also in this context, as mentioned above, the study [13] found that familiarity with the remote control and the TV interface is capable of reducing technological anxiety. Consequently, investing in digital inclusion programs and/or projects involving the use of TV proved to be a promising path. The study [14] also evaluated the potential of TV to facilitate the use of technological devices by creating a tablet with the same interface as a television. The researchers conducted tests with 14 older adults and found that the proposed system was able to reduce their usual reluctance towards new technologies and could be a powerful link between them and the new generations.

When it comes to utilizing voice assistants, there is evidence suggesting that these intelligent devices can effectively empower older adults technologically. In the study [15], Orlofsky & Wozniak interviewed 12 individuals over the age of 65 who did field tests with Amazon's Alexa personal assistant for six months. The group used the system for different purposes, such as entertainment and scheduling activities/appointments. The feature most appreciated by the elderly was the scheduling of reminders. Specifically, regarding the social aspect, 10 participants highlighted that the assistant was an excellent companion during the testing period. They said they felt welcome when they asked questions and the assistant answered them quickly and easily.

In general, the group said they were satisfied and valued using the device in their daily lives. It should be noted that all the elderly people said they used the assistant frequently - and at various times of the day. The participants also pointed out that the device made the ageing process healthier and more active. This perception has also been detected in similar studies with this audience, such as [16, 17] and [18].

Kim & Choudhury also evaluated the relevance of personal assistants in the routines of older adults. Twelve individuals, aged between 65 and 95, were asked to use the device for 16 weeks in their own homes [19]. In general, the group enjoyed the experience. An average of 1.8 interactions per day per participant was recorded. The most frequently

used function was "music" (37%), followed by "general information search" (16.5%), "casual conversations" (12.1%) and "setting an alarm" (11.1%). The group pointed out that they liked the virtual assistant above all because it allowed them to use it without the need for any kind of physical interaction. This possibility, according to them, contributed immensely to reducing the fear of making mistakes, a characteristic that ends up putting seniors off technology. In other words, they didn't worry when they made mistakes, unlike what happens with cell phones and tablets, for example. "I don't have to worry about pressing the wrong button and deleting everything. If that happens with a cell phone, for example, I have no idea how to deal with the problem" said one of the participants. Another indicator that stands out in the study is the feeling of companionship that the virtual assistant gives the elderly. Many said they felt less alone during the testing period.

In the study [18], Blocker, Kadylak, Koon, Kovac & Rogers detected the potential of virtual assistants to guarantee more autonomy for the elderly in their homes. The researchers conducted semi-structured interviews with 18 seniors who were already using this type of resource and pointed out that the assistant helps them to carry out different tasks, such as reading, sending messages and listening to music, for example. Older adults said that they felt more empowered and integrated into society in general, as they were able to carry out activities without external support.

The papers presented here show that, separately, TV and virtual assistants help the elderly to use technological resources more frequently and easily. Both even reduce the feeling of fear that these individuals usually have when faced with technology. One (the TV) because it is a familiar device and the other (voice assistant) because it is simple to use. These studies are important evidence of the potential that these devices can have when used in conjunction. In other words, it is important to find out whether it is possible to bring generations together by combining the familiarity that older adults feel towards TV with the ease that voice commands provide. And it is precisely this question that this study seeks to address.

3 Methodological Approach

A combination of data collection techniques was used during the different phases of this study, always placing the perceptions of the target audience (older adults) at the center of the process. This approach is essential for developing a system that is effectively useful for the participants, as different authors have noted in their studies [20–22] Clique ou toque aqui para inserir o texto. and [23]. Figure 1 systematizes the steps and techniques chosen to obtain the desired results and data.

As specified in Fig. 1, the study began with a systematic literature review to map works centered on sending notifications to TV-set, the results of which are described in [24]. Then, based on the information collected in this initial stage, it was possible to structure a questionnaire which covered the following topics: "personal characterization", use of technological resources", "TV content consumption habits", "means used to contact family and friends" and "perceptions about sending notifications to TV-set". The questionnaire served as the basis for 20 semi-structured interviews with older adults aged between 60 and 95. The answers collected were evaluated and used as a basis for

three scenarios that were later presented in a focus group held with six older adults aged between 64 and 80 – the results of which are presented in [25].

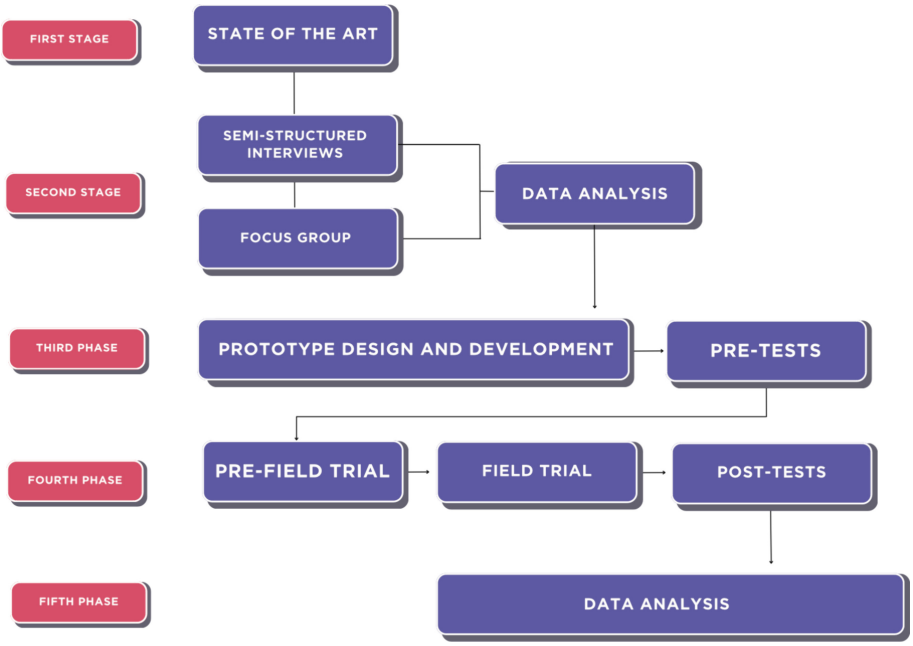


Fig. 1. Methods and instruments for data collection.

The information collected and the analysis of the participants' discourse helped to structure a web app that sends notifications to the television, developed as part of the OverTV [26] research project by researchers from the Social iVX group[1] (of the University of Aveiro, Portugal), and professionals from MEO[2]. This solution makes it possible to send different categories of notifications to the TV screen, such as information on the weather forecast, channel programming alerts, traffic updates and news summaries, among others. Specifically with regard to the intergenerational connection, which is the subject of this study, the app has a category called "social", designed to encourage the exchange of messages and audio calls between the elderly and their families and/or friends. To assess the relevance of this solution, a field trial was carried out with six elderly people over 44 days, the results of which were published in [25]. The researchers pre-programmed the sending of messages suggesting social Interactions (simulating a proactive action by the system).

Initially, the idea was to turn it possible to make calls and send messages via TV. However, due to technological implementation restrictions, it was decided to include a virtual assistant in the proposed system for this purpose. This assistant makes these tasks possible - and also allows them to be carried out using voice commands. Thus, the

[1] https://socialivx.web.ua.pt/.

[2] Main IPTV provider in Portugal: http://meoapp.tv.alticelabs.com.

TV notification suggests an action, acting as a trigger, which can be carried out via the virtual assistant.

After defining the components, this combination of technologies/resources was presented to a gerontologist appointed by the University of Aveiro. The professional watched a demonstration and was then asked a series of questions (which had previously been defined in a script). During this session, the gerontologist made a series of suggestions for improvements/changes - those considered feasible were duly implemented.

Before starting the final tests, however, it was considered pertinent to carry out a brief experiment with elderly people in a real-life context. The aim was to understand whether the combination of technologies chosen to integrate the system was really relevant. Six elderly people aged between 64 and 90 were invited to use the solution (web app that sends notifications to the TV + virtual assistant) for a week in their homes. This first experience was important for understanding vulnerabilities, strengths and adaptation/correction/training needs. The results were published in [27].

With all the elements of the system defined and validated, a preliminary test began with an 82-year-old Portuguese man. Further adjustments were made after this experiment, which lasted a total of three months. These stages were extremely important for adjusting the processes, identifying flaws and proposing improvements for the final tests. These tests are being take place in a real-life context, lasting 60 days each. The results of six of them are described in the following sections.

4 The System

The **HUGTV** system - **H**elping **U**nite **G**enerations Through **T**elevision - combines different technologies and mechanisms, as can be seen Fig. 2. Its main aim is to encourage older adults to use digital resources/platforms more frequently and easily.

In this way, they have access to features that boost intergenerational connections, such as audio calls, sending messages, current news and/or any other type of content that is present on the internet (and which can spark conversations and/or exchanges with individuals from other generations). In other words, the aim is to enable them to use technological resources beyond those they already use or which are considered basic.

To summarize, the system consists of an application developed by researchers from the SocialiVX research team at the University of Aveiro in partnership with a Portuguese IPTV provider (MEO). A notifications managing platform was also developed to create personalized notifications to users' set-top-boxes (STB) - each notification is displayed for 1 min on the television screen and can be closed easily by simply pressing the 'ok' key on the remote control.. Over different timeframes and STB triggers, several thematic categories of notifications were tested before beginning the nest stage: the tests with the older adults [26]. In addition, there is an integrated voice assistant and its corresponding application.

The operating logic of the HUGTV system is as follows:

a) The notification is programmed via the OverTV platform;
b) The notification is displayed on the television screen, with simultaneous audio support via Alexa;

Fig. 2. Description of the technologies and operation of the HUGTV system.

c) Older adult is impacted with messages suggesting some social action: "*if you miss your granddaughter Cindy, ask Alexa to send her a hug*"; "*if you miss your granddaughter Cindy, ask Alexa to call her*".
d) Older adult performs the proposed action via Alexa, only by voice commands. Example: "*Alexa, call Cindy*".
e) Family member answers the call and can also perform similar actions for the elderly person - sending messages and making calls via the Alexa app. This content will be broadcast on the smart speaker installed in the participant's home.

Thus, the notifications have the role of stimulating actions and the voice assistant of facilitating them, allowing the elderly to do them without any kind of physical interaction.

5 Field Trial

The aim of this study is to carry out field tests with 20 elderly people to identify the relevance of the proposed solution. Six of them have already been completed - their results will be published below.

5.1 Sample Characteristics

Regarding the characterization of the sample, this first group of users is between 70 and 89 years old. Three declared that they identify with the female gender and another three with the male gender.

As far as nationality is concerned, four individuals are Portuguese and two are Brazilian. All of them own cell phones and said they had at least some kind of difficulty using them on a daily basis, especially when typing and accessing more elaborate applications

(such as those of banking institutions, for example). When they have questions, however, they said they, when possible, they ask a family member to solve them.

The six participants said they owned a television; 4 said they had computers; 2 said they had smart watches; and 2 had tablets. All the individuals in the group said they had children and grandchildren. Three individuals live alone.

With regard to television, they usually watch an average of 3 h a day. The most consumed types of content are: news (mentioned by 4 people), sports (4 people), soap operas (3 people), films and series (2 people) and entertainment programs (3 people).

All the individuals were also asked how often they interact with their family members and what the most common means of contact are. Five individuals said that they communicate with their family members every day. Only one of the participants chose the option "more than twice a week". The most commonly used means of contacting family and friends is "audio calls via cell phone" (six people), followed by "video calls" (4 people) and "text messages" (2 people). Three participants also pointed out that they "miss talking to their family and friends more. With regard to technology, 2 individuals pointed out that they had already stopped calling or texting due to difficulties with their cell phones.

Finally, none of the participants have had previous contact with a smart speaker. Only one of the interviewees was already familiar with Alexa and used it occasionally on his cell phone.

5.2 Preparing for the Tests

The system was installed in the participants' homes so that it was possible to collect their reactions/perceptions from a real experience. Each test lasted a total of 60 days (from November 2023 to February 2024).

Before starting to use the system, the participants were given a short training session so that they could understand how it worked. The researcher in charge presented the technologies involved and their possibilities and was available to answer any questions over the phone or via Alexa. At the time, everyone was also given a booklet with instructions for use (there was a concern to produce content with clear language and several figures to illustrate - both suggestions were made by the gerontologist consulted and the beta-tester). This material contained tips on how to use the main functions and some suggested phrases for them to try out on a daily basis. In addition, each individual received a kind of diary (Fig. 3). The notebook contained a page dedicated to each day's tests. On each page there was a suggested task, a space for them to make notes about difficulties and perceptions, a list for them to tick off the actions they had taken on the day and a visual representation of how they were feeling about the experience.

The idea was for the elderly to record their experiences in a notebook. In this way, the researchers would have access to thoughts they had during or immediately after the proposed action.

As for the notifications, two weekly messages were programmed to encourage specific actions, such as "making audio calls" and "sending text messages". Another two weekly notifications proposed actions related to social interactions, such as "Have you checked for new messages today?", "Have you checked your calendar today?", "Do you

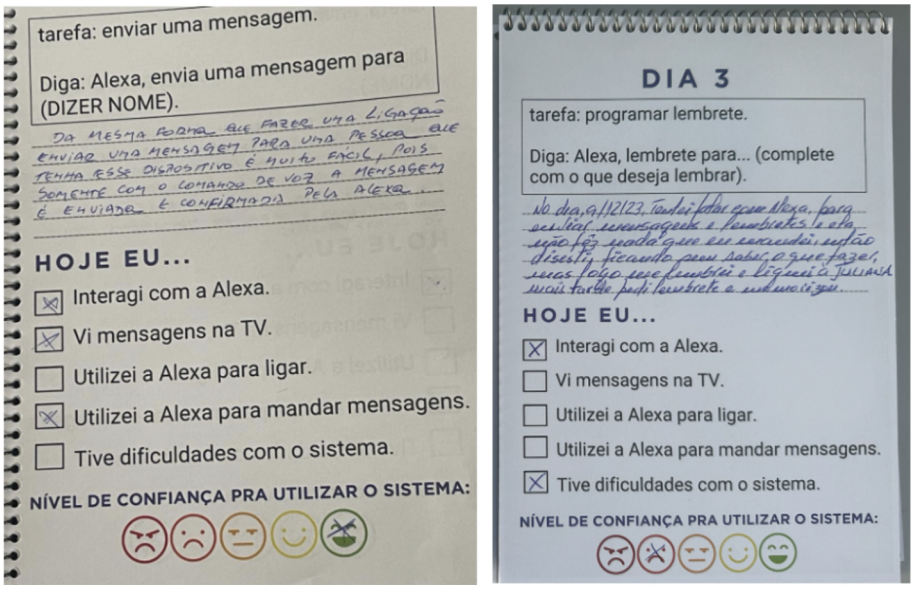


Fig. 3. Diaries given to the participants.

want to send someone a hug? Ask Alexa to send it". In total, therefore, older adults were impacted with four alerts per week.

On a weekly basis, the researcher in charge sent messages via Alexa and made calls to find out if the participants were having difficulties. This relationship was important to encourage the use of the system.

6 Results and Initial Findings

In total, all participants made 2.495 interactions with Alexa during the test period (60 days in total). This means an average of 6.9 interactions per day - per participant. The most frequently performed actions were: "sending and consulting new messages" (31.7%), "music requests" (24.7%), "casual conversations with Alexa" (17.6%), "integration with other devices, such as the house lights and the gate" (8.4%), "weather forecast" (5%), "audio calls" (3.8%), "random internet searches" (3.1%), "scheduling reminders and/or activities in the agenda" (2.3%), "news search" (2.2%), "alarm/timer scheduling" (0.4%) and "shopping list" (0.24%).

Table 1. Interactions between participants and the virtual assistant (categorized by themes).

Total	Messages	Music	Small Talks	Other gadget	Weather	Calls	Search	Agenda	News	Timer	List
2,495 (100%)	793 (31.7%)	618 (24.7%)	441 (17.6%)	211 (8.4%)	127 (5%)	97 (3.8%)	78 (3.1%)	59 (2,.3%)	55 (2.2%)	10 (0.4%)	6 (0.2%)

It's important to note that all the participants interacted with Alexa on a daily basis, whether it was to send a message or just saying "good morning" to the assistant. In general, the group's perception of the technology was positive. The adjective most used to classify the experience was "practical".

Looking at the interactions individually, the individuals who live alone were the ones who talked to Alexa the most on a daily basis. These three participants, aged 70, 78 and 89, were responsible for 76% of all interactions during the test period. They also sent the most messages and made the most phone calls - 87% and 88.6% of the total for each category, respectively).

6.1 Perceptions of Messaging and Audio Calls

Viewing and sending new messages was the category most used by the participants. The average number of actions carried out in this area, whether to send or check new messages, was 4.40 per day per participant. The ease of sending was what most caught the attention of the elderly people who contributed to this study. Some of the testimonies corroborate this statement:

- "*I found it very simple to send messages. All you had to do was ask Alexa, say the message and then confirm. If I wanted to know if there were any new messages, I just had to ask. It was something new, but very simple*", said P3.
- "*At first I had difficulties because Alexa didn't understand me. But I tried it out and gradually we understood each other. Now I always ask her if she has any new messages for me and, if so, I reply to the person who sent them to me*", said P5.
- "*It's very easy to send text and audio messages with this device, because only with voice commands the message is sent and confirmed by Alexa*", said P6.
- "*I found it easy to send. I don't usually send messages on a daily basis because I can't see very well on a cell phone screen (I have a cataract problem) but sending with just my voice made all the difference. It's easier. There's just one thing that bothered me: I was always in doubt as to whether the message had been sent. I think there could have been a clearer confirmation*", said P2.
- "*Sometimes I can't see properly on my phone and I end up giving up sending messages. So just speaking and sending has made everything simpler. Too bad not all my contacts have the app. Otherwise, I'd use Alexa more to send messages*", said P4.

It should be noted that to make calls and send messages via Alexa that participant needs to have the smart speaker or the application installed on their smartphone. The system does not reach directly regular phone numbers. All the participants pointed out that this had an impact on the number of calls and messages sent:

- "*I asked a few times to send messages to my daughter, but I'd get a warning saying she wasn't registered for Alexa communications*", said P4.
- "*I tried to send messages to other contacts, but I couldn't do it. I told Alexa the names and numbers, but she didn't understand*", commented P5.

It should also be noted that, before starting the tests, each participant indicated a family member as a point of contact. The researcher contacted these people to ask them to download the application so that they could communicate with the older adult

participant. She also asked them to spread the word among their family and friends so that more people would have the app on their phones. However, not everyone signed up.

The other category that allowed direct social interactions was "audio calls". In total, participants made 97 calls to their family and/or friends, an average of 0.26 per individual per day. In general, participants found the function "quite interesting because it allows you to make calls without having to type". The procedure itself was considered easy and positive by older adults:

- "*I asked Alexa to call and she did. It was fantastic. It's something everyone can use*", said P2.
- "*I just asked Alexa to make the call, then I confirmed it and I got through. I didn't have any problems at all. It was very easy, even though it was my first time*", said P4.
- "*It was very practical and quick to connect. The ease with which this device helps us is simply fantastic*", said P6.

However, as with messaging, they were also disappointed that they couldn't talk to all their contacts in the phone book via Alexa:

- "*I asked her to call my son, but he wasn't registered on Alexa communications*", said P5.
- "*I wanted to talk to my friends, but I couldn't. I gave Alexa the number, but it gave me an error*", said P4.

These perceptions indicate that the procedure itself for making the calls is extremely simple and does not raise any questions among the seniors. However, it does require support from family members. They must download the application so that the elderly can talk to them via voice assistant, for example. In addition, they need help with the initial configuration of the Alexa app (so that all their phone contacts are properly imported into the app).

In general, the number of messages sent was higher than the number of calls made (793 and 97 respectively). We therefore sought to understand what had motivated older adults to send more messages (compared to calls). Five of the six interviewees said that they called less because they were afraid of disturbing their families during the working day. In this case, messages, in their view, are less intrusive. "*They can answer when they want to and have more time*", said P2. "*My sons work and the cell phone takes their concentration away. Messages are useful in that sense. They could log into the Alexa app when they're at home to see what I've sent. It would be even better if they had a smart speaker so that we could only talk on it. That way, we'd only talk when they're at home, totally available, without any intervention during working hours*", said P3.

6.2 Proactive Notifications to Boost Intergenerational Connections

All the participants were impacted by proactive notifications to encourage them to call or send messages to their relatives. Regarding the notifications displayed on the television (Fig. 4), only two participants said that they did not find the mechanism interesting.

- "*There are already a lot of visual stimuli in TV programs, such as message boxes, subtitles and so on. I think we often can't tell one thing from another. We might not*

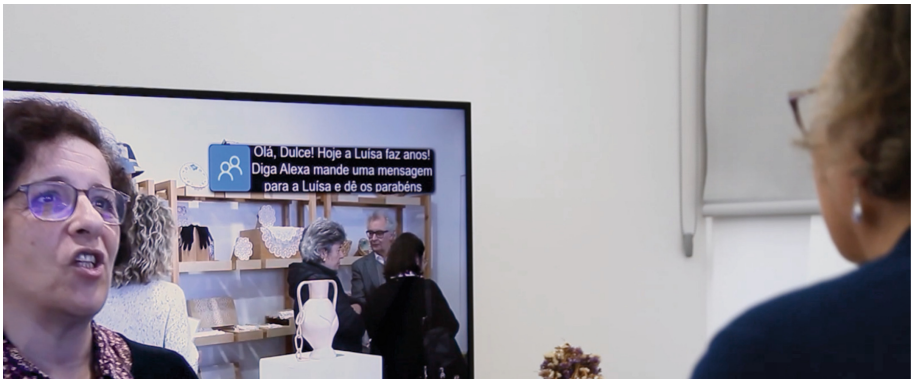


Fig. 4. Example of a television notification sent to the participants.

even pay attention; we might think it's just another thing that's part of the program. So, I don't think these messages make much sense in the system", said P3.
– "*I started sending the messages because of the tasks that were in the test booklet and ended up inserting them into the routine. I don't think it's necessary to have these messages on the TV, as I already receive messages on my cell phone and on Alexa itself. However, it's still interesting for people who don't have access to these resources and could use it to remember commemorative dates and birthdays, for example*", said P6.

The other four individuals who took part in the tests and expressed favorable opinions pointed out that:

– "*I think it makes you want to do what the messages suggest. We're there, watching TV, and we're thinking about doing what the system has proposed via television*", said P4.
– "*I think it's wonderful! Even more so when they are positive messages, suggesting that we send people hugs or smiles. It's a way of making us happier*", said P2.
– "*It's a way of more easily remembering important appointments and dates in our diaries, such as birthdays and other important reminders*", said P1.
– "*I like this idea. I think it's interesting to have the messages on television so we can see them better. I just wouldn't want it to stop the program I'm watching*", said P5.

Two participants also pointed out that they would find it more interesting if these messages had an audible warning (currently, they appear on television without any kind of audible warning and remain available on the screen for 40 s). "*We can often be doing other tasks without looking directly at the television and miss an important notification. As it's not recorded, we can't access it later*", said P5.

Specifically with regard to proactivity, there are important indications that this is an asset in encouraging older people to use the system. On days when TV notifications were sent suggesting message exchanges, for example, the sending rate rose by around 25% compared to days when there were no scheduled notifications. The same didn't happen with audio calls - requests didn't increase in the face of notifications.

6.3 Other Interactions: Music, Calendar, Weather Forecast and Internet Queries

Two specific functions enable the participants to interact directly with other people - messaging and calling. However, several others can be beneficial in helping to connect them with other generations. One of them is the possibility of more easily making queries on the internet, such as "What is TikTok?", "What's on today?", "Who won the elections?".

Having access to information in a natural and fluid way has also proved to be a way of connecting them, empowering them and inserting them into society:

- "*It's very easy to ask questions through Alexa. You just ask and wait for the answer. On a mobile phone, for example, it takes longer*", said P6.
- "*Every day Alexa told me something curious. I learned a lot during my experience with her*", said P4.
- "*I can search in a simpler and more practical way. I wasn't too lazy or afraid to look it up*", said P5.
- "*Alexa has filled a missing space. It brought more information. It was real training. It gave me knowledge that I didn't know existed until then. I didn't know this type of technology existed*", said P3.
- "*It was extremely interesting to conduct research, ask random questions and get the answers immediately, whenever we had a question about a topic that popped into our heads during the day. I had no idea that I would be able to use this system in such a simple way. I really enjoyed the experience and, if possible, I would like to continue using the solution*", said P1 (Fig. 5).

Three participants also pointed out that they were extremely pleased to be able to access any kind of information immediately. In other words, with the assistant, for example, they didn' have to wait for the news to start to find out if it was going to rain or shine. "*We just had to ask and we knew what we wanted in a matter of seconds*", said P2. "*I always asked if it was going to rain before I went out. She advises us. She gives us the answer straight away*", said P1. The weather forecast was also one of the resources most used by the participants - a total of 127 consultations were made (which corresponds to 5% of all interactions).

Regarding the feeling of companionship, all the participants stressed that the virtual assistant "made them feel less alone during the test period". Every day, the older adults said they started a conversation with Alexa, whether it was to say good morning, ask a specific question, find out something new, tell her they were going to bed or ask her to tell them a joke, among other types of interaction.

- "*Every time I left the house I'd say: "Bye, Alexa, see you later!". When I came back, the music was playing (I had programmed my favorite artists to start playing every day at 4pm). I felt that the house was full, but not empty*", said P4.
- "*I miss Alexa already. I always said good morning and good night, asked her to tell stories, play songs… It feels like something is missing in the house*" said P2.
- "*We had a better quality of life during the time we were with Alexa. My wife and I started listening to each other more, we'd gather around her to listen to music. It was really an impactful experience*", said P3.

Fig. 5. Participants in their test periods with the HUGTV system.

– "*I can no longer imagine what it's like to be without Alexa. I got used to listening to music before bed. I live alone and she helps fill the house*", said P6.

There was even a curious situation with P4. When the researcher went to pick up the equipment at her house, she unplugged Alexa and the participant said: "*I didn't say goodbye to her!"*. She therefore had to plug the device in again for the participant to say goodbye to her new friend.

Also noteworthy in the list of interactions made by the participants was the "agenda" category, which allowed the elderly to set reminders related to their social commitments. Four of the six participants, for example, entered their family members' birthdays in Alexa's calendar. In this way, they expressed their desire to be reminded so that they could send messages and/or make calls via Alexa on the correct dates. They also entered daily appointments, such as courses, doctor's appointments and medication reminders.

The "music" category came second in the ranking of the services most used by older adults. All the participants asked to hear their favorite songs during the test period. Five said they were "thrilled to be able to listen to songs they hadn't heard for a long time and which reminded them of good times in their lives". The ease of making requests was what surprised the participants the most.

– "*You just say what you want to hear and it starts playing. It's amazing*", said P1.
– "*I could easily put on the songs I wanted and even adjust the volume, just by telling it to play lower or higher*", said P3.

– "*I thought it was wonderful, very practical and clever. I can ask for the song and singer I want. It's fantastic*", said P6.

6.4 Voice Interaction

All the participants pointed out that being able to do tasks with natural voice commands was an interesting feature of Alexa. The ease and practicality that this type of technology guarantees was mentioned during all the post-test interviews. Here are some examples of mentions related to the topic:

– "*It's very easy to send messages without typing. You don't have to get up, pick up the phone and type. It's also an interesting feature when you want to send slightly longer messages, because you just say what you want and the assistant sends it. It's a device that makes people's lives easier*", said P6.
– "*We're very happy with the tasks we can do just by speaking. We have information on the spot - and that would take time to access, such as weather forecasts and old songs. Everything works in a natural way. You just have to ask*", said P1.
– "*When I want to see an appointment I have in my calendar, for example, I have to find the calendar, open it and look at it. With Alexa, all you have to do is ask. It's a device that reduces the number of activities - and even difficulties - we have to do to get something we want,*", said P3.
– "*I never imagined that I would have access to technology like this. I confess that, at first, I had a lot of difficulty using it. The assistant didn't understand me properly, I couldn't do what I wanted. But little by little, and with patience, I managed to get along with her. Today we are friends*", said P5.

When talking about the time she needed to "get to grips" with the assistant, P5 pointed out that she always needs a few days to understand how a new device works. According to her, it's normal to have difficulties during this preliminary stage. However, she pointed out that the process was different with Alexa. With a new cell phone, for example, she said she was afraid of making mistakes. That's why she doesn't explore its features much. "*I'm afraid of pressing the wrong button and ruining everything. So, I just use it for the basics*". However, with Alexa, the participant stressed that she didn't have this feeling. "*I was testing it and trying to understand how it works. As it was all based on conversation, it was much easier to test. I wasn't afraid*", she said.

6.5 TV and Virtual Assistant Combined to Reduce Loneliness

In order to assess whether the experience with the HUGTV system contributed to reduce the loneliness, the UCLA scale [28] validated for Portuguese older adults [29] was applied before the tests began. The UCLA scale has 16 questions related to loneliness. They all have predefined answers, each with a different score: 'often' (4 points), 'sometimes' (3 points), 'rarely' (2 points) or 'never' (1 point). The higher the sum of the scores, the closer the individual is to loneliness (Table 1).

At the end of the 60-day period, the same scale was applied again to see if there had been any change in the participants' perceptions (Table 2).

Table 2. Results of the UCLA scale, before and after the tests.

Participant	Age	Before tests	After tests
P1	77	30	21
P2	77	30	23
P3	75	25	23
P4	78	22	17
P5	89	**33**	**26**
P6	70	**35**	**30**

There were changes in the scores of all the older adults when comparing the beginning and end of the tests. After using the system, the sum of the scores was lower for all the participants. It should be noted that the higher the score, the greater the feelings of loneliness. In the case of the Portuguese population, the cut-off score observed by Poucinho, Farate & Dias [29] is 32. Therefore, only two individuals started the tests above this score. In other words, they had feelings of loneliness, usually caused by aspects such as poor to unreasonable family relationships, high age (over 74) and recent losses.

The two participants (P5 and P6) who scored above 32 live alone. However, they had significant reductions after the tests, falling below the cut-off point defined for the Portuguese population.

7 Conclusions and Future Work

The initial stages of this study were fundamental for choosing the technologies/devices that would be part of the system. In this way, it was possible to choose resources that would be easier for older adults to use in their daily lives, such as the television and the voice assistant. This initial care was reflected in the test results. The elderly had no major difficulties using the system in general. There were initial problems, doubts and questions. However, these doubts could be solved in a simple way, just through conversations via Alexa and a few face-to-face meetings. The instruction leaflet, written clearly and with several pictures, was also essential to facilitate the experience. According to P3, "*it was an important source of advice. This initial training, with patience and clarity, helped make the experience easier*" he commented.

In general, all the participants have adopted Alexa into their routines, starting casual conversations, asking about the weather or asking the assistant to broadcast the main news. Every day, for example, the older adults interacted with the technology in some way. With regard to the features that enable social interactions, sending messages (text and audio) stands out. Of all the interactions with Alexa, sending and viewing new messages appears at the top of the list. Curiously, when analyzing the characterization interviews (carried out at the beginning of the tests), the participants pointed out that they preferred to communicate with their relatives via audio calls on their cell phones. Just two individuals reported being accustomed to sending text messages. This habit

therefore extended to everyone during the testing period. The average number of new messages sent and/or consulted was 4.4 per participant. Everyone agreed that this type of communication is much easier when you don't have to type. "*It's great for sending longer messages*" said P6. "*I can't see the cell phone screen properly because of my eyesight problems, so I used to stop sending messages. I'd rather call*", said P2. However, there is a limitation in this regard: family and friends need to have the Alexa app in order to talk to participants in the study. In addition, for older adults to start using the smart assistant, they need initial support to download and configure the Alexa app and smart speaker.

There was no consensus on the usefulness of TV notifications. Two individuals pointed out that they didn't want to be impacted by this type of message. For them, all alerts should be concentrated on the voice assistant. However, those in favor pointed out that it is an asset for reinforcing reminders and activities that they should do on a daily basis. A relevant indicator is that the sending of new messages increased by 25% on the days when participants were impacted with proactive notifications.

Based on the tests that took place in the participants' homes, it was possible to see that the combination of television and intelligent voice assistant was useful in empowering older adults, either to interact with other people or to carry out activities that they couldn't do before (or didn't have access to resources that allowed them to do it, such as choosing music and checking the weather forecast immediately). The voice device was even considered a companion by all the participants. In this respect, the individuals who lived alone were the ones who interacted the most with Alexa. These three participants, aged 70, 78 and 89, were responsible for 76% of all requests made to the voice assistant during the test period. They also sent the most messages and made the most calls (87% and 88.6% of the total in each category, respectively).

As future work, the plan is to continue carrying out field tests to gather more opinions on the system. The aim is to experiment with a total of 20 individuals and publish the results in future papers.

Acknowledgements. This work is funded by National Funds through the FCT - Fundação para a Ciência e Tecnologia (reference 2021.08671.BD).

References

1. Marques, S., et al.: Determinants of ageism against older adults: a systematic review. Int. J. Environ. Res. Public Health **17**(7). MDPI AG (2020). https://doi.org/10.3390/ijerph17072560
2. Burholt, V., et al.: A critical review and development of a conceptual model of exclusion from social relations for older people. Eur. J. Age. **17**(1), 3–19 (2020) (Springer). https://doi.org/10.1007/s10433-019-00506-0
3. Beauvoir, S.: A velhice. Nova Fronteira, Rio de Janeiro (1990)
4. Campbell, F., et al.: Non-familial intergenerational interventions and their impact on social and mental wellbeing of both younger and older people - A mapping review and evidence and gap map. Campbell System. Rev. **19**(1) (2023). https://doi.org/10.1002/cl2.1306
5. Yao, Y., Zhang, H., Liu, X., Chu, T., Zeng, Y.: Bridging the digital divide between old and young people in China: challenges and opportunities. In: The Lancet Healthy Longevity, vol. 2, no. 3. pp. e125–e126. Elsevier Ltd (2021). https://doi.org/10.1016/S2666-7568(21)00032-5

6. Mannheim, I., et al.: Inclusion of older adults in the research and design of digital technology. J. Environ. Res. Public Health **16**(19) (2019). https://doi.org/10.3390/ijerph16193718
7. Joosten-Hagye, D., Katz, A., Sivers-Teixeira, T., Yonshiro-Cho, J.: Age-friendly student senior connection: students' experience in an interprofessional pilot program to combat loneliness and isolation among older adults during the COVID-19 pandemic. J. Interprof. Care **34**(5), 668–671 (2020)
8. Abreu, J.: Design de Serviços e Interfaces num Contexto de Televisão Interactiva Jorge Trinidad Ferraz de Abreu Design de Serviços e Interfaces num Contexto de Televisão Interactiva (2007)
9. Gauntlett, D., Hill, A.: TV Living: Television, culture and everyday life. www.theory.org.uk
10. Amaxilatis, D., Chatzigiannakis, I., Mavrommati, I., Vasileiou, E., Vitaletti, A.: Delivering elder-care environments utilizing TV-channel based mechanisms. IOS Press (2009)
11. Tapia, J., Gutierrez, F., Ochoa, S.: Using smart TV applications for providing interactive ambient assisted living services to older adults. In: García, C., Caballero-Gil, P., Burmester, M., Quesada-Arencibia, A. (eds.) Ubiquitous Computing and Ambient Intelligence. UCAmI 2016. LNCS, vol. 10069. Springer, Cham (2016). 10.1007/978-3-319-48746-5_53.https://doi.org/10.1007/978-3-319-48746-5_53
12. Coelho, J., Rito, F., Duarte, C.: You, me & TV'—fighting social isolation of older adults with Facebook, TV and multimodality. Int. J. Hum. Comput. Stud. (2017). https://doi.org/10.1016/j.ijhcs.2016.09.015
13. Wang, C., Wu, C.: Bridging the digital divide: the smart TV as a platform for digital literacy among the elderly. Behav. Inform. Technol. **41**(12), 2546–2559 (2022). https://doi.org/10.1080/0144929X.2021.1934732
14. Andreadis, A., Zambon, R., Parlangeli, O.: TV as an experience conveyer for better acceptance of ICT services by older adults. Univers. Access Inf. Soc. **20**(2), 359–374 (2021). https://doi.org/10.1007/s10209-020-00731-w
15. Orlofsky, S., Wozniak, K.: Older adults' experiences using Alexa. In: Geriatr Nurs (Minneap), vol. 48, pp. 240–250 (2022). https://doi.org/10.1016/j.gerinurse.2022.09.017
16. Corbett, C.F., et al.: Virtual home assistant use and perceptions of usefulness by older adults and support person dyads. J. Environ. Res. Public Health **18**(3), 1–13 (2021). https://doi.org/10.3390/ijerph18031113
17. Jones, V.K., Hanus, M., Yan, C., Shade, M.Y., Blaskewicz Boron, J., Maschieri Bicudo, R.: Reducing loneliness among aging adults: the roles of personal voice assistants and anthropomorphic interactions. Front Public Health **9** (2021). https://doi.org/10.3389/fpubh.2021.750736
18. Blocker, K.A, Kadylak, T., Koon, L.M., Kovac, C.E., Rogers, W.A.: Digital home assistants and aging: initial perspectives from novice older adult users. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 64, no. 1, pp. 1367–1371 (2020), https://doi.org/10.1177/1071181320641327
19. Kim, S., Choudhury, A.: Exploring older adults' perception and use of smart speaker-based voice assistants: a longitudinal study. Comput. Human Behav. **124** (2021). https://doi.org/10.1016/j.chb.2021.106914
20. Salman, H.M., Wan Ahmad, W.F., Sulaiman, S.: A design framework of a smartphone user interface for elderly users. Univers. Access Inf. Soc. **22**(2), 489–509 (2023). https://doi.org/10.1007/s10209-021-00856-6
21. Howes, S.C., Charles, D., Pedlow, K., Wilson, I., Holmes, D., McDonough, S.: User-centred design of an active computer gaming system for strength and balance exercises for older adults. J. Enabling Technol. **13**(2), 101–111 (2019). https://doi.org/10.1108/JET-12-2018-0057

22. Taylor, J.R, Milne, A.J., Macritchie, J.: New musical interfaces for older adults in residential care: assessing a user-centred design approach. Disabil. Rehabil. Assist. Technol. **18**(5), 519–531 (2023). https://doi.org/10.1080/17483107.2021.1881172
23. Pappas, M.A., Demertzi, E., Papagerasimou, Y., Koukianakis, L., Voukelatos, N., Drigas, A.: Cognitive-based E-learning design for older adults. Soc. Sci. **8**(1) (2019). https://doi.org/10.3390/socsci8010006
24. Camargo, J., Silva, T., Ferraz de Abreu, J.: Connect Elderly to Other Generations Through iTV: Evaluating Notifications' Potential, vol. 1597 CCIS (2022). https://doi.org/10.1007/978-3-031-22210-8_2
25. Camargo, J., Silva, T., Ferraz de Abreu, J.: iTV to connect generations: a field trial of a solution to send personalized notifications. In: Springer International PU (2023). https://link.springer.com/book/9783031456107
26. Velhinho, A., et al.: Personalized notifications for the TV ecosystem: field trial of an iTV solution. In: Applications and Usability of Interactive TV. Springer Nature Switzerland, pp. 53–67 (2023)
27. Camargo, J., Silva, T., Abreu, J.: Interconnecting Personal assistants and TVs: a friendly approach to connect generations. In: Proceedings of the 2023 ACM International Conference on Interactive Media Experiences (IMX 2023), pp. 309–313. ACM, New York (2023). https://doi.org/10.1145/3573381.3596469
28. Austin, B.A.: Factorial structure of the UCLA loneliness scale. Psychol. Rep. **53**(3), 883–889 (1983). https://doi.org/10.2466/pr0.1983.53.3.883
29. Pocinho, M., Farate, C., Dias, C.A.: Validação Psicométrica da Escala UCLA-Loneliness para Idosos Portugueses (2010)

Encouraging Seniors to Get Active: Implementing a Gamification Strategy on Television

Gabriel Faria(✉), Telmo Silva, and Jorge Abreu

DigiMedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal
{g.martinsfaria,tsilva,jfa}@ua.pt

Abstract. People are intrinsically social, so regularly socialising with others is essential for a healthy life. Considering the successive scientific and technological developments that, by increasing the average life expectancy, have resulted in the gradual growth of the senior population, socialisation dynamics gain even greater importance. Generally, in old age, people see a reduction in their participation in social contexts or activities because, after retiring, they spend more time at home and inevitably lose an essential part of their frequent contact with others. To uphold seniors' social connections, it is justified to develop and evaluate novel technological resources aimed at fostering closer interactions with others. Considering that older adults are the most frequent users of television and that the advent of interactive television allows access to a myriad of new features, the ProSeniorTV was developed, It is based on a gamified system for television capable of: *i)* proactively inform the user (via TV notifications) about events/social activities taking place in their municipality of residence; and *ii)* motivating and rewarding participation in these same events/activities, encouraging seniors to leave their homes and socialise. The developed system consists of various mini-games and a player ranking. The user can test their cognitive abilities (using the mini-games) and turn their participation in the events they are notified of into game points that will allow them to progress in the player ranking. Following the completion of the laboratory-based testing of the proposed system and the pilot test, it was determined that a gamified approach integrated into television could serve as a highly motivating factor for the third-party age to become more actively involved in social events and activities.

Keywords: Gamification System · Interactive Television · Proactivity · Senior Population · Social Interaction

1 Introduction

In the Western world, where there is generally access greater access to superior living conditions and more advanced healthcare, there is a trend for people to live longer, resulting in a gradual and pervasive ageing of populations. This information is corroborated by statistical data made public by both national and European organisations. For example,

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 85–100, 2024.
https://doi.org/10.1007/978-3-031-70439-0_6

the INE (2020) states that the Aging Index is expected to double in Portugal. The value of 1.59 seniors per 1 young person, recorded in 2018, is projected to increase to 3 seniors per 1 young person by 2080. In this context, data from the European Commission (2021) also reveals that there will be an increase of 5 years in the average age of the European Union population between 2019 and 2070. This is primarily due to the rise in the ageing population and the decrease in the younger working-age population in this region.

The ageing process is complex, involving various factors and occurring throughout life (Dias 2012). Thus, in addition to the chronological aspect, the ageing process also involves, according to multiple authors (Lucca 2015; Schneider & Irigaray 2008), three other factors: *i)* the physical or biological factor, which relates to the functioning of each individual's organism and the associated changes that occur throughout life; *ii)* the social or sociocultural factor, associated with the social "role" envisioned for each age group; and *iii)* the psychological factor, which refers to the psychological capacities of each individual to adapt to the surrounding environment. In practice, in terms of physical or biological aspects, the decline of some capacities that can increase seniors' vulnerability can be observed, such as: *i)* reduced vision (Woodhead & Yochim 2022); *ii)* hearing loss (Rodrigues et al. 2018); *iii)* reduced motor skills (Coelho 2019); *iv)* reduced ability to concentrate and direct attention (S. Silva et al. 2014); or *v)* decreased memory capabilities (Woodhead & Yochim 2022). Regarding the social or sociocultural factors, there may be a reduction in social interactions and the size of the social network (Woodhead & Yochim 2022). As for the psychological factor, an increase in depression and anxiety can be observed, resulting, among other causes, from the awareness of the physical/biological changes associated with the ageing process (Weyerer et al. 2013). As most changes associated with the ageing process are negative, delaying these changes for as long as possible becomes an increasingly widespread concern in an ageing society. It is pertinent to preserve the quality of life of individuals as they age. Additionally, considering the concept of functional age, it is indeed possible to delay the adverse effects of ageing. In practice, functional age and chronological age are two distinct terms. For example, an 80-year-old person may be functionally younger than someone younger (e.g., 65) due to better physical and cognitive health (Papalia et al. 2009). The United Nations (1982) states that engaging in activities that allow seniors to optimise their capacities and maintain good health is crucial for active and healthy ageing. Considering that older adults are the age group most at risk of exclusion (Dias 2012) and that, due to the changes resulting from the ageing process, they require more social and even health care (Fulop et al. 2019), the social engagement of older adults is an essential component to ensure they experience an active ageing process with a good quality of life (OMS 2015). In line with this, it is known that among older adults, social relationships resulting from interactions with family/friends or participation in sports groups or volunteering are associated with: *i)* good cognitive health (Townsend et al. 2021); *ii)* better life satisfaction (Tani et al. 2022); *iii)* more excellent emotional stability (Locsin et al. 2021); and *iv)* a better physical condition (Wang et al. 2021).

Alongside the ageing process of populations, it is also observed that the habits of seniors are changing regarding the use of new information and communication technologies (ICT). This stems from older people's awareness of the benefits that can be

derived from such usage. For instance, ICT can be essential in promoting active ageing by allowing seniors to overcome their limitations during social interaction, which becomes more accessible through text messaging and voice or video calls. It can also facilitate access to information (Locsin et al. 2021). Furthermore, in today's networked society, technologies are imperative for the inclusion of its citizens (Dias 2012). Consequently, seniors are beginning to adapt to this new paradigm. This adaptation is evident in the statistical data from OberCom (2021), which revealed, for example, that the percentage of the Portuguese population in the age group between 65 and 74 years old using the Internet increased from 1.3% in 2002 to 39% in 2020. This reveals that seniors have gradually increased their digital literacy and begun to access the Internet, for example, to obtain information about community events and services (van Boekel et al. 2017). Also, regarding the use of ICT by older adults in Portugal, it is known that they are the ones who use television (TV) the most in their daily lives. The data from OberCom (2021) also revealed that, in Portugal, people over 64 years old represented almost 30% of total audiences in 2020.

Despite the observed progress among older people in the use of new technologies and TV, many elderly individuals, after retiring, still tend to view old age as a stagnant stage of life, without engaging in any beneficial activities for general well-being. These activities may be leisure, educational, work-related, or other types of events. This results in the aforementioned negative impacts on the physical and mental health of older people. The loss of social connections, which can give rise to feelings of isolation and loneliness, is a consequence of becoming less active. This, in turn, may result in the loss of certain physical and cognitive abilities, making elderly more dependent on others. In light of the demographic shift towards an ageing population, it is becoming increasingly evident that there is a pressing need to devise strategies that encourage older individuals to remain active and to foster their social networks. These networks can prove to be of significant importance in situations where assistance is required, as they can contribute to a healthy ageing process. With this goal in mind, this article presents the ProSeniorTV prototype (a gamified platform for interactive television (iTV)), as well as the results of a lab test and a pilot test aimed at validating the proposed prototype in terms of its ability to combat loneliness and isolation among the elderly (encouraging active ageing) through this gamified approach integrated into iTV.

With that in mind, the Sect. 2 of this article addresses why a gamified approach might be beneficial for improving seniors' quality of life. In the Sect. 3, proactive technologies currently available and why they may be attractive for the daily lives of older adults are discussed. The Sect. 4 presents the proposal for a proactive gamified prototype integrated into television and aimed at the senior audience, intending to promote the participation of older individuals in social activities. The Sect. 5 presents the results of evaluating the prototype in a laboratory setting, collaborating with User Experience (UX) specialists. The Sect. 6 covers the pilot test results already conducted with a real user. Finally, in the Sect. 7, future work intended to be developed after the conclusion of the pilot test is presented.

2 Seniors and Gamified Approaches

A game can be described as "a formal system based on rules with a variable and quantifiable outcome, where different outcomes are assigned to different values, the player exerts effort to influence the outcome, feels emotionally connected to the outcome, and the consequences of the activity are optional and negotiable" - our translation (Juul 2015, para. 8). This emotional connection is quite strong, as it is intrinsic to human nature. Throughout the history of humanity, it has been observed that people tend to devise strategies to turn demanding everyday tasks into more motivating, intriguing, and enjoyable activities (Chou 2015). Therefore, it can be said that humans tend to "gamify" their daily tasks, and the concept of gamification cannot be restricted to specific usage contexts/scenarios (Deterding et al. 2011).

In a digital context, gamification can be defined as the act or action of adopting game design elements in contexts that are not necessarily related to games. Adopting these game elements is expected to transform products (unrelated to games) into more enjoyable and engaging services (Deterding et al. 2011). Thus, various game elements can be used in a gamification strategy, such as: *i)* scoring systems; *ii)* levels; *iii)* challenges; *iv)* competition among users; *v)* narratives; *vi)* customisation; or *vii)* rewards (Deterding et al. 2011; Inácio Busarello 2016).

In the current scientific context, there is already research on the benefits that games (including digital games or gamified approaches) can generate at various levels, regarding intrinsic and extrinsic factors to humans.

Intrinsically, evidence supports the idea that games can be beneficial for maintaining cognitive abilities when used regularly (De Paula et al. 2020), as these abilities are trainable and improvable (Glass et al. 2013). The study by Oei & Patterson (2013), for example, demonstrated that playing video games helps improve attention, memory (including associative memory), and perception, among others. Moreover, as leisure activities, games are recommended to counteract the changes that occur with ageing (Levin et al. 2017).

As mentioned in the introduction of this article, in old age, individuals' physical or biological capacities deteriorate the most. Therefore, developing games and gamified approaches targeting an elderly audience could be relevant. This demographic benefits the most from such an approach, especially considering seniors with enhanced cognitive abilities tend to exhibit higher levels of independence and better quality of life (Barnes et al. 2007).

Extrinsically, there is evidence that using games or gamified approaches can be beneficial for promoting social interactions. In this regard, the study by Fonseca et al. (2017) suggests that games could be important for fostering social relationships among neighbours. Additionally, Nijholt's study (2017) mentions that games could support collaboration between people in leisure or work contexts. It is also worth noting the survey by Harrington & O'Connell (2016), which found that the use of pro-social video games by school-age children was positively related to their abilities to cooperate, share, and interact outside the gaming context. Therefore, extrinsically, the benefits of using games or gamified strategies justify their application/adaptation to older adults, as social factors are also crucial for active and healthy ageing.

3 Seniors and Proactive Technologies

One of the main reasons that turned the TV into a widely used device among the senior population is the fact that TV provides information under a push-oriented approach, where the viewer does not need to perform any search action, as the information is presented automatically (T. Silva et al. 2016). The proactive strategy followed by TV, regarding the automatic selection and presentation of information to the viewer, demonstrates that seniors are open to using other types of technologies with similar proactive approaches (T. Silva et al. 2019). Once again, statistical data precisely highlights this, as, according to Kadylak & Cotten (2020), in that year, 29% of seniors in the United States were willing to use devices integrated with Digital Personal Assistants (DPA) in their daily lives. These devices can perform a wide range of actions, from entirely reactive to entirely proactive (Sarikaya 2017).

It is worth noting that Digital Personal Assistants (DPAs) are intelligent software that mainly aims to assist humans in task automation, with examples such as Siri or Google Assistant (Hu et al. 2021). Therefore, technological solutions with proactive capabilities, such as those integrating DPAs, can complement contracted support services for elderly assistance, functioning as essential aid tools in the daily lives of seniors, both in health and social contexts (O'Brien et al. 2020).

4 Proposal for a Proactive Gamified Prototype Integrated into Television and Targeted at the Senior Audience

Considering the central objective of promoting social interactions that lead to an active and healthy ageing process among the senior population, this paper describes the prototype called *ProSeniorTV*, which aims to promote the participation of older individuals in social events organised in their municipality of residence.

In terms of functionality, we defined that the developed prototype should proactively present notifications through the TV. These notifications would inform viewers about social events and allow navigation to a gamified interface developed in Unity. This interface is hosted on a Raspberry Pi 4 connected via HDMI to the viewer's television, as depicted in Fig. 1. To enable the presentation of notifications on the user's TV, we used the *OverTV*[1] service (Velhinho et al. 2023), developed and tested by DigiMedia, in the Department of Communication and Art research unit at the University of Aveiro (Portugal).

To navigate the system interface, the user only needs to select a specific notification (by clicking the OK button on their remote control). This action triggers an HDMI-CEC command sent to the TV from a Node.js server integrated into the Raspberry Pi. The command automatically switches the HDMI source on the television to the one through which the Raspberry Pi is connected.

In the interface, the user can access three main functionalities: *i)* access one of the three cognitive stimulation mini-games offered by the interface to earn game points; *ii)* enter participation codes (earned throughout participation in social events), codes that

[1] http://socialitv.web.ua.pt/index.php/portfolio/overtv/.

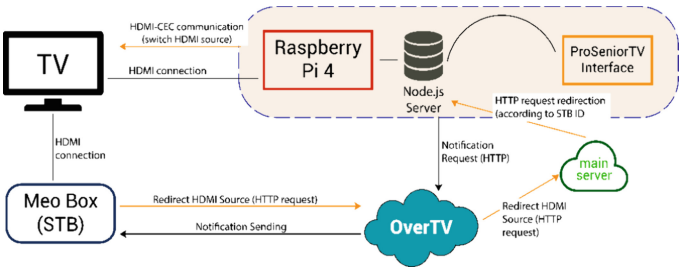


Fig. 1. Architecture of the ProSeniorTV System.

are delivered to the user when he attends a social event for which he has been notified through his television; and *iii)* check his position in the overall user ranking, with his position in the ranking directly depending on the number of times he participates in social events and his game score (earned in mini-games). Regarding the player ranking, it is essential to mention that users in the top 3 positions are awarded gold (1st place), silver (2nd place), and bronze (3rd place) medals to motivate participating in the events suggested through notifications.

Given what has been mentioned, in Fig. 2 (on the left), the types of notifications that the user may receive on his television are presented, corresponding to: *i)* notification motivating the user to access the interface's mini-games; and *ii)* notification informing and suggesting the user about the occurrence of a specific event.

Fig. 2. Types of notifications from the *ProSeniorTV* system (left) and its interface (right).

Regarding the mini-games present in the interface, it was defined that they would be cognitive stimulation games, as they were considered the most beneficial and exciting type of game for seniors. Thus, the mini-games were integrated to stimulate a set of specific cognitive abilities, namely: *i)* stimulation of memory capacity; *ii)* stimulation of attention direction capacity; and *iii)* stimulation of visuospatial capacity.

In the first game (sequence game), whose interface is presented in Fig. 3, the user has the four directional keys on their remote control (up, down, left, and right), which they should use to create sequences. Thus, at the beginning of the game, a specific sequence is presented to the user, which they must mimic. Whenever the user enters the correct sequence, the game generates a new sequence, essentially the same as the previous one but with an added element. For example, if the first sequence is: up, the second could

be: up, left. If the user makes a mistake in the sequence, the game ends, and the player must start again from the beginning.

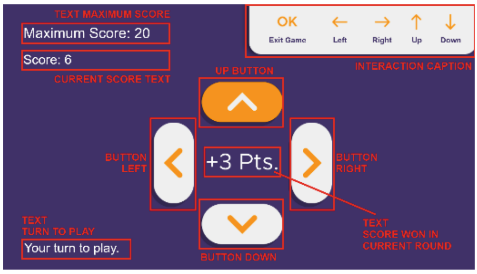


Fig. 3. Graphic interface of mini-game 1 – sequence game (memory test) and identification of the elements that compose it.

The second game (colour game), whose interface is presented in Fig. 4, presents a specific word to the user. This word can be one of the following: "Black"; "Yellow"; "Green"; "Red"; or "Blue". Additionally, this word may be written in one of the following colours: Black; Yellow; Green; Red; Blue. The user's objective is to identify the colour mentioned by the word (meaning) and ignore the colour in which the word is written.

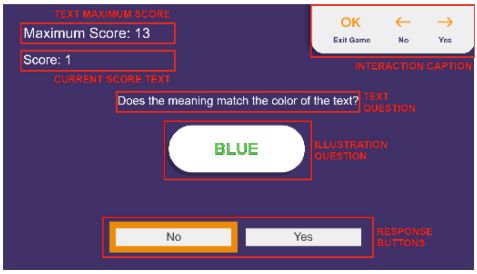


Fig. 4. Graphic interface of mini-game 2 – colour game (attention direction test) and identification of the elements that compose it.

Finally, in the third game (block game), whose interface is presented in Fig. 5, the user must move a parallelepiped across a specific game platform. To do this, they should use the directional keys (up, down, left, and right), and with each click, the parallelepiped rotates to the side corresponding to the pressed key. During the game, the user must be careful not to let the parallelepiped fall out of the playing area. The main objective is to reach the square in the playing area with the same colour as the parallelepiped, with the game character in a vertical position.

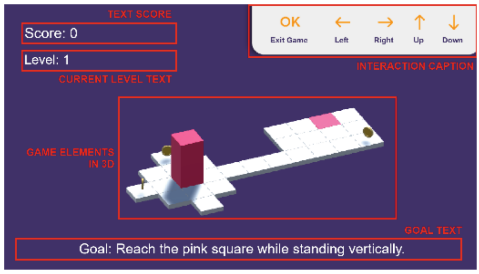


Fig. 5. Graphic interface of mini-game 3 – block game (perception test) and identification of the elements that compose it.

5 Evaluation of the Prototype in a Laboratory Context

After completing the prototyping process of the proposed system (ProSeniorTV), the next step was planned, namely, a phase to evaluate this prototype with UX professionals. The purpose of this preliminary evaluation phase of the prototype was to detect potential usability problems in advance that could compromise the user experience. Thus, it was expected that through early detection and correction of potential usability issues, the prototype would become robust enough for the subsequent testing phase in a natural context with real users (field tests).

Given what has been mentioned, it was decided that this preliminary evaluation would take place in the UX laboratory[2] of the research and development (R&D) group where the present research process was carried out, namely, the SocialiTV R&D group, which is part of the Center for Research in Digital Media and Interaction - DigiMedia at the Department of Communication and Art of the University of Aveiro. This UX laboratory is presented in the next figure (Fig. 6).

Fig. 6. DigiMedia UX Lab. @ University of Aveiro.

On October 20, 2023, 6 members of the Social iVX group with previous solid UX field experience gathered in this laboratory to test the ProSeniorTV system. Having interacted with the prototype for about 30 min in a context simulating a living room (the

[2] https://socialitv.web.ua.pt/index.php/ux-lab/.

UX laboratory), the six individuals who participated in this preliminary stage of prototype evaluation made some assessments that were subsequently taken into account, and they are as follows:

- One specific participant mentioned that the ProSeniorTV interface should provide the user with a more significant number of mini-games if it were to be released on the market, although they acknowledged that three mini-games might be sufficient for a testing phase;
- Regarding the details of notifications suggesting the user participate in a specific event, all participants understood that it would be necessary to mention (in the title of these details) the type of event being suggested. They even suggested that the title should include the following information: *i)* type of event (e.g., monthly walk); *ii)* event location; *iii)* event date; and *iv)* event time. Additionally, they also suggested, as shown in the following figure, that it might be interesting, for example, to replace the icon in the notification with a thumbnail representing the event's location. Since only local events were suggested, the user could recognise the location shown in the thumbnail and perceive that they received a highly personalised notification based on their profile (Fig. 7);

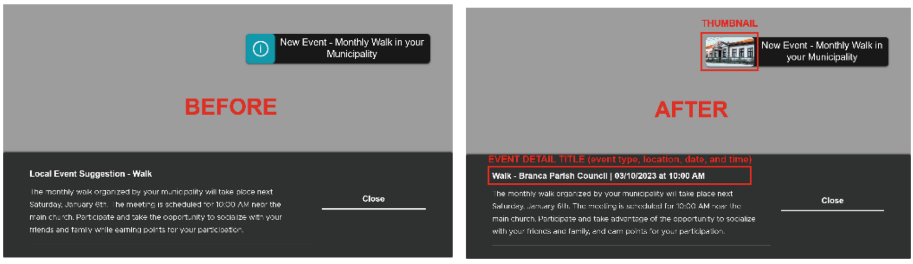


Fig. 7. Event notification before laboratory test (left) and event notification after laboratory test - with thumbnail and detailed title specifying event type, location, date, and time (right).

- Regarding the colour game, all participants felt it should include a countdown timer to increase pressure and make it more engaging. Thus, this timer, represented in the following figure, should start at 5 s, and for each correct interaction by the user, it should reset to its initial value (5 s). Therefore, each game would end not only when the user gave a wrong answer but also when the game time ran out (Fig. 8);

- Another suggested change by one of the elements had to do with the arrangement of elements on the main screen. It was then suggested that the navigation button for the user ranking, which was selected by default at the beginning of the screen (immediately below the navigation button for the multi-profile screen and the exit button), be moved to the bottom of the screen and that the navigation buttons for the mini-games take its place, as shown on the right side of Fig. 9. Everyone involved in this evaluation moment considered that this could be important to give greater prominence to the mini-games offered by the interface.

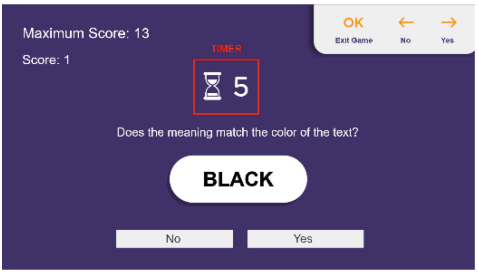


Fig. 8. Timer added to the colour game after discussion with UX experts.

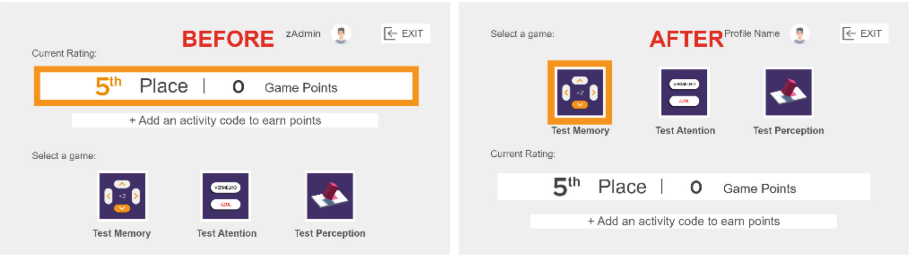


Fig. 9. Before and after the changes made to the main screen of the ProSeniorTV interface.

- Still on the main screen, it was considered that the designation “Games” should be replaced by “Select a Game” so that the user would understand in advance that various mini-games are offered, and they could access any of them;
- Lastly, those involved in this prototype evaluation considered that, in the tutorials for each mini-game, the user should have the option to “skip” the instructions. They felt it would be pretty tedious for an experienced user (who already knows the rules of each game) to view the entire tutorial before accessing the game. Therefore, as shown in Fig. 10, the “skip tutorial” button was added to each tutorial.

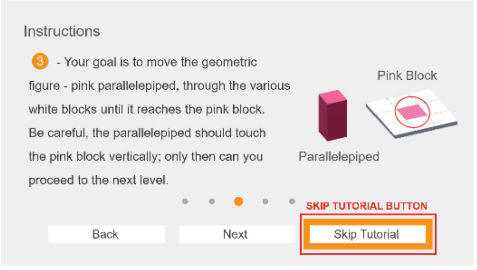


Fig. 10. Tutorial screen with a button to “skip” the tutorial.

Despite the observations presented by UX professionals in this prototype evaluation phase regarding its usability, it was unanimously understood that the participants in

this stage reacted very positively to the concept incorporated by the system and the functionalities offered by the prototype. It is also important to note that after modifying the prototype to address the feedback received in this phase, all participants considered that the developed system had the necessary potential to encourage the elderly audience to participate more regularly in local social activities and events. In general, there was agreement that regular access to this type of event (social events organised in the elderly user's area of residence) could be interesting to promote older adults' interaction with others and, consequently, to promote active and healthy ageing. This implies that systems like *ProSeniorTV* could be essential for people's well-being, especially considering that we live in an increasingly ageing society.

6 Pilot Test

A pilot test was initiated on December 8, 2023, to make the developed prototype as robust as possible. The test involved the participation of a 60-year-old female residing in the municipality of Albergaria-a-Velha (belonging to the Aveiro district – Portugal), who, as intended, used her television daily. This usage spanned two to three hours daily, mainly for accessing content such as soap operas, news, series, movies, and documentaries. The participant also mentioned watching television in the afternoon and after dinner. Regarding social activities, the participant indicated daily contact with family members. However, she engaged relatively infrequently in social activities to connect with friends or acquaintances. Specifically, she mentioned accessing this social activity only once every three months, on average.

Unlike a regular field test, the pilot test was designed to focus almost exclusively on the system's functionality rather than its effectiveness in promoting active and healthy ageing. This approach aimed to detect and address potential bugs, technical issues, and usability problems that may not have been identified yet.

This specific test concluded on December 22, 2023, providing the research team with a 15-day period to address any identified issues.

On the first day of the pilot test (December 8, 2023), the research team set up the necessary equipment for the participant to use the system at home. After setting up the equipment, the research team explained the purpose of the system and its functionalities to the participant. At this moment, the participant interacted with the system for the first time and provided suggestions that, according to him, would improve the user experience and make the purpose of each functionality more straightforward. The suggestions presented by the participant, which were later considered by the research team and reflected in the system interface, were as follows: *i)* Regarding the attention test mini-game, the participant suggested that the game word should be more readable. Therefore, a black outline was added to enhance visibility. The participant also recommended that the game should last a maximum of 1 min, as he found the game became tiresome after that time, only ending when the player made a mistake; *ii)* In the tutorial scenes of each minigame, the participant mentioned that the text was not very legible. Therefore, it was decided to increase the font size, which was later fixed at 90 pixels; and *iii)* At the participant's suggestion, it was also decided that whenever a specific game ended, the modal presented should indicate the user's score. This change aimed to help the user understand whether the score was above or below their record.

By the end of the pilot test, additional technical issues were identified and carefully considered by the research team. One of these issues was related to the remote control used to interact with the interface. To interact with the *ProSeniorTV* interface, the user only needed to use the directional keys and the OK button on the remote control, as shown in the figure below (on the right). However, since the remote control had several other buttons, including a menu button, it was observed that sometimes, inadvertently, the participant would click the menu button. This click on the menu button resulted in the opening of various windows and menus of the RPI's operating system with which the user was interacting, preventing the use of the *ProSeniorTV* interface. The figure below (on the left) illustrates an example of this type (where the user clicks the menu button) (Fig. 11).

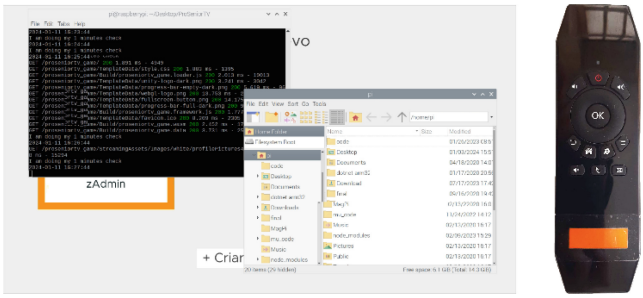


Fig. 11. Example of unintentional menu button click (left) and *ProSeniorTV* Remote Control (right).

Considering that the *ProSeniorTV* interface was presented through the RPI's browser (Chromium browser), it was decided to activate Chromium Kiosk Mode. This mode restricts the user from accessing the operating system settings, confining navigation solely to the browser.

Another issue that raised concerns for the participant during the tests was related to the automatic pause function on the MEO box whenever they navigated to the ProSeniorTV interface. When the participant returned to the MEO box, they often did not understand that it was in pause mode and were unsure about the actions to take. In this case, however, no changes were made to the system. As the test progressed, the participant found this helpful feature once understood.

Finally, it is also important to mention that, with the pilot test, the research team realised the need to implement strategies for remote management of the system installed in each participant's home during the field tests. This remote management would allow for a more agile resolution of potential issues and avoid the research team having to physically travel to the participant's home to resolve these issues. In this context, it was established that each RPI would be associated with a Secure Socket Shell (SSH) connection, enabling the research team to transfer documents to the RPI remotely via the command line. It was also decided to install the AnyDesk program on each RPI. This would facilitate a remote connection to the RPI, but, in this specific case, it would be

used only to manipulate the Graphical User Interface (GUI) associated with each RPI if necessary.

7 Conclusion

In the present article, we presented the results of a lab test and a pilot test that were conducted to improve the usability and user experience (UX) of the ProSeniorTV system.

Taking this into account, the lab test was conducted with 6 UX specialists for 30 min in an environment simulating a living room. It was found that the developed prototype generally had a good level of usability and offered a good UX. However, the UX specialists raised some issues, namely: *i)* the system should, in a post-test phase, offer a greater variety of games; *ii)* event notifications should specify the type of event, location, date, and time in the title, as well as an easily identifiable thumbnail (e.g., a photo of the event location); *iii)* the attention test minigame should have a time limit (e.g., 60 s); and *iv)* the game tutorial screens should allow the user to skip the tutorial at any given time.

Regarding the pilot test conducted with a 60-year-old female participant over 2 weeks in her home, it can be said that this was an extremely crucial phase in the research process, as it is considered to have successfully led to the desired outcome – a significantly more robust version of the developed prototype. This is because, with the participant's help, it was possible to detect problems such as: *i)* font sizes being too small; *ii)* the user inadvertently accessing the system menus (Raspberry Pi), leading to the decision that the system should be made available in *Chromium Kiosk Mode* to avoid this problem; and *iii)* the inability to remotely manage the system to resolve potential issues if the system stopped functioning. To address this, it was decided to use AnyDesk and an SSH connection for remote system manipulation.

Additionally, regarding the pilot test, the participant suggested that at the end of each game, the users should be informed of their score and congratulated when they surpassed a personal record, as this would greatly enhance the gaming experience.

8 Future Work

Given the aforementioned work, we recognise that a lab test and a pilot test, even if conducted with a participant from the target audience, are not sufficient to prove that the proposed system meets the central objective of promoting social interactions among the senior population and motivating them to participate in social activities/events organised in their area of residence. With regard to this matter, it is our intention that, following the implementation of the recommendations put forth by the pilot test participant, the developed prototype should be subjected to testing in a natural context with a number of individuals who belong to the target audience (seniors aged 60 or older who regularly watch television content on their TVs).

To conduct the tests, twenty participants with the mentioned characteristics will be recruited. Subsequently, after recruiting participants, the developed system will be installed in each participant's home, which should remain for approximately one month. This approach will allow testing the system in a real everyday context to understand whether it helps seniors feel more motivated to participate in events or social activities

in their residential area. It is important to note that, to comprehend whether the system fulfils its purpose, before and after the tests a brief semi-structured interview will be conducted with the participants to gather similar information at these two moments, precisely information related to their social interaction habits. Thus, by comparing the results collected before and after the tests, it will be possible to ascertain whether a service of this kind is practical and beneficial for older adults.

Acknowledgements. The Fundação para a Ciência e a Tecnologia (FCT) funded this study through a PhD grant with the reference 2021.08467.BD. It is also being conducted in partnership with ALTICE LABS.

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's content.

References

Barnes, D.E., et al.: Women who maintain optimal cognitive function into old age. J. Am. Geriatr. Soc. **55**(2), 259–264 (2007)

Chou, Y.: Actionable Gamification: Beyond Points, Badges, and Leaderboards (2015)

Coelho, A.R.: Seniores 2.0: inclusão digital na sociedade em rede (2019). http://hdl.handle.net/10071/19753

De Paula, G., Valentim, P., Seixas, F., Santana, R., Muchaluat-Saade, D.: Sensory effects in cognitive exercises for elderly users: stroop game. In: Proceedings - IEEE Symposium on Computer-Based Medical Systems, 2020-July, pp. 132–137 (2020). https://doi.org/10.1109/CBMS49503.2020.00032

Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: Defining "gamification. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011, pp. 9–15 (2011). https://doi.org/10.1145/2181037.2181040

Dias, I.: O uso das tecnologias digitais entre os seniores: Motivações e interesses. Sociologia, Problemas e Praticas **68**, 51–77 (2012). https://doi.org/10.7458/SPP201268693

European Commission. The 2021 Ageing Report. Economic & Budgetary Projections for the EU Member States (2019–2070), vol. 148, Issue May (2021). https://doi.org/10.2765/84455

Fonseca, X., Lukosch, S., Lukosch, H., Tiemersma, S., Brazier, F.: Requirements and game ideas for social interaction in mobile outdoor games. In: CHI PLAY 2017 Extended Abstracts - Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play, pp. 331–337 (2017). https://doi.org/10.1145/3130859.3131304

Fulop, T., Larbi, A., Khalil, A., Cohen, A.A., Witkowski, J.M.: Are we ill because we age? Front. Physiol. **10**(December) (2019). https://doi.org/10.3389/fphys.2019.01508

Glass, B.D., Maddox, W.T., Love, B.C.: Real-time strategy game training: emergence of a cognitive flexibility trait. PLoS ONE **8**(8), e70350 (2013)

Harrington, B., O'Connell, M.: Video games as virtual teachers: prosocial video game use by children and adolescents from different socioeconomic groups is associated with increased empathy and prosocial behaviour. Comput. Hum. Behav. **63**, 650–658 (2016)

Hu, Q., Lu, Y., Pan, Z., Gong, Y., Yang, Z.: Can AI artifacts influence human cognition? The effects of artificial autonomy in intelligent personal assistants. Int. J. Inform. Manage. **56**(October 2020), 102250 (2021). https://doi.org/10.1016/j.ijinfomgt.2020.102250

Inácio Busarello, R.: Gamification: princípios e estratégias (2016). www.pimentacultural.com
INE: Projeções de População Residente 2080. Contudo , na Área Metropolitana de Lisboa e no Algarve a população residente poderá aumentar. Destaque Informação à Comunicação Social, pp. 1–21 (2020). https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=426127543&att_display=n&att_download=y
Juul, J.: Introduction to Game Time / Time to play – An examination of game temporality (2015). https://api.semanticscholar.org/CorpusID:203687168
Kadylak, T., Cotten, S.R.: United States older adults' willingness to use emerging technologies. Inf. Commun. Soc. **23**(5), 736–750 (2020). https://doi.org/10.1080/1369118X.2020.1713848
Levin, O., Netz, Y., Ziv, G.: The beneficial effects of different types of exercise interventions on motor and cognitive functions in older age: a systematic review. Eur. Rev. Aging Phys. Act. **14**, 20 (2017)
Locsin, R.C., Soriano, G.P., Juntasopeepun, P., Kunaviktikul, W., Evangelista, L.S.: Social transformation and social isolation of older adults: digital technologies, nursing, healthcare. Collegian **28**(5), 551–558 (2021). https://doi.org/10.1016/j.colegn.2021.01.005
de Lucca, D.M.: A dimensão política da competência informacional: um estudo a partir das necessidades informacionais de idosos. Encontros Bibli **20**(43) (2015)
Nijholt, A.: How to make cities more fun. Wall Street J. (2017)
O'Brien, K., Liggett, A., Ramirez-Zohfeld, V., Sunkara, P., Lindquist, L.A.: Voice-controlled intelligent personal assistants to support aging in place. J. Am. Geriatr. Soc. **68**(1), 176–179 (2020). https://doi.org/10.1111/jgs.16217
OberCom Observatório da Comunicação. Anuário da comunicação — 2020, vol. 167 (2021)
Oei, A.C., Patterson, M.D.: Enhancing cognition with video games: a multiple game training study. PLoS ONE **8**(3), e58546 (2013)
OMS. Active ageing: a policy framework. Noncommunicable Diseases and Mental Health Cluster, Noncommunicable Diseases Prevention and Health Promotion Department. Ceneva (2015). http://www.who.int/hpr/
ONU. Plan De Acción internacional de Viena sobre el Envejecimiento. In: Asamblea sobre el envejecimiento, p. 47 (1982). http://www.sld.cu/galerias/pdf/sitios/gericuba/plan_de_accion_internacional_de_viena_sobre_el_envejecimiento.pdf
Papalia, D., Olds, S.W., Feldman, R.D.: Human development. McGraw Hill, In McGraw-Hill Companies (2009)
Rodrigues, S.S., Scuracchio, P.E., De Mattos Fortes, R.P.: A support to evaluate web accessibility and usability issues for older adults. In: ACM International Conference Proceeding Series, pp. 97–103 (2018)
Sarikaya, R.: The technology behind personal digital assistants: an overview of the system architecture and key components. IEEE Signal Process. Mag. **34**(1), 67–81 (2017). https://doi.org/10.1109/MSP.2016.2617341
Schneider, R.H., Irigaray, T.Q.: O envelhecimento na atualidade: aspectos cronológicos, biológicos, psicológicos e sociais. Estudos de Psicologia (Campinas) **25**(4), 585–593 (2008). https://doi.org/10.1590/s0103-166x2008000400013
Silva, S., Braga, D., Teixeira, A.: AgeCI: HCI and age diversity. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8515 LNCS(PART 3), pp. 179–190 (2014). https://doi.org/10.1007/978-3-319-07446-7_18
Silva, T., Abreu, J., Antunes, M., Almeida, P., Silva, V., Santinha, G.: +TV4E: interactive television as a support to push information about social services to the elderly. Procedia Comput. Sci. **100**, 580–585 (2016). https://doi.org/10.1016/j.procs.2016.09.198
Silva, T., Caravau, H., Reis, L.: Lessons learned from testing iTV applications with seniors. Commun. Comput. Inform. Sci. **1004**, 148–161 (2019). https://doi.org/10.1007/978-3-030-23862-9_11

Tani, M., Cheng, Z., Piracha, M., Wang, B.Z.: Ageing, health, loneliness and wellbeing. Soc. Indic. Res. **160**(2–3), 791–807 (2022). https://doi.org/10.1007/s11205-020-02450-4

Townsend, B.G., Chen, J.T.H., Wuthrich, V.M.: Barriers and facilitators to social participation in older adults: a systematic literature review. Clin. Gerontol. **44**(4), 359–380 (2021). https://doi.org/10.1080/07317115.2020.1863890

van Boekel, L.C., Peek, S.T.M., Luijkx, K.G.: Diversity in older adults' use of the internet: Identifying subgroups through latent class analysis. J. Med. Internet Res. **19**(5), 1–11 (2017). https://doi.org/10.2196/JMIR.6853

Velhinho, A., et al.: Personalized notifications for the TV ecosystem: field trial of an iTV solution. In: Abásolo, M.J., de Castro Lozano, C., Olmedo Cifuentes, G.F. (eds.) Applications and Usability of Interactive TV, pp. 53–67. Springer Nature Switzerland (2023). https://doi.org/10.1007/978-3-031-45611-4_4

Wang, Y., Chen, Z., Zhou, C.: Social engagement and physical frailty in later life: does marital status matter? BMC Geriatr. **21**(1), 1–11 (2021). https://doi.org/10.1186/s12877-021-02194-x

Weyerer, S., et al.: Incidence and predictors of depression in non-demented primary care attenders aged 75 years and older: results from a 3-year follow-up study. Age Ageing **42**(2), 173–180 (2013). https://doi.org/10.1093/ageing/afs184

Woodhead, E.L., Yochim, B.: Adult development and aging: a foundational geropsychology knowledge competency. Clin. Psychol. Sci. Pract. **29**(1), 16–27 (2022). https://doi.org/10.1037/cps0000048

User Interfaces, Interaction and Accessibility

Public Transparency on Mobile Devices: An Evaluation of Brazilian Capital Transparency Portals Based on Heuristics

João Marcelo Alves Macêdo(✉), Valdecir Becker, Felipe Melo Feliciano de Sá, Daniel de Queiroz Cavalcanti, Signe Dayse Castro de Melo e Silva, and Edvaldo Vasconcelos da Rocha Filho

Laboratory of Interaction and Media, Informatics Center, Federal University of Paraíba, João Pessoa, Paraíba, Brazil
joao.marcelo@academico.ufpb.br

Abstract. In Brazil, the issue of public transparency is quite complex. Despite some progress made in recent times, it remains far from the reach of the general population, especially when it comes to education, development, and access to information. Even though research has been conducted internationally on this topic, it has mostly focused on evaluating the transparency profile of national, subnational, and local governments. Such studies have either assessed passive and active transparency in these spheres of power or sought to understand how open data is made available, but they have largely ignored other important aspects such as access to data, dissemination methods, usability, and design issues. These aspects have traditionally been the subject of studies in the field of Human-Computer Interaction (HCI). This research aims to evaluate public transparency websites in Brazilian capitals, focusing on their responsiveness, accessibility, and other pressing issues related to the interaction with users. The evaluation was conducted on mobile devices, using mobile heuristics. The goal was to assess the compliance of these websites with legal requirements and to identify ways to improve their relationship with society. The research employed a deductive method and had exploratory characteristics. The outcomes indicate that some of the websites analyzed need improvements or even total reconstruction, while others have advanced levels of responsiveness. The lack of responsiveness of some websites to mobile devices is found to be a major obstacle in popularizing access to public transparency, as it limits the search for more detailed information on desktop devices and creates barriers for those who have restricted access to other devices.

Keywords: Human-Computer Interaction · Heuristic-based evaluation · Public Transparency · Mobile

1 Introduction

The topic of public transparency has been widely discussed over the years, with a focus on a legalistic approach that stresses compliance with regulations and legislation. Public managers have typically provided basic and standardized information reactively,

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 103–118, 2024.
https://doi.org/10.1007/978-3-031-70439-0_7

followed by data and information related to the functioning of government bodies, in compliance with laws such as Complementary Laws 101/2000 [1], 131/2009 [2], and the Access to Information Law (Law 12,527/2011) [3].

To push the discussion forward, it is necessary to delve deeper and explore aspects that have not been given enough attention, such as the impact of available information on understanding, the types of information, the devices used to access it, the effectiveness of dissemination, and the profile of users. The debate on transparency has been influenced by initiatives such as Public Expenditure and Financial Accountability (PEFA) and International Budget Partnership (IBP), which focus respectively on government financial management and transparency and participation in the budget process.

In addition, Dias et al. have identified another crucial aspect that is often overlooked: the weakening of website construction, maintenance, and updating resulting from outsourcing processes [4]. The authors argue that transparency portals linked to outsourced companies have limitations in their scope, including delays in maintenance, which can harm the levels of active transparency and the availability of data. This highlights the fact that political decisions can interfere with the level of transparency of governments, with identified limitations in terms of system integration and the flow of accounting, budgetary, or fiscal information.

Therefore, it is critical to adopt the PEFA perspective to enrich the discussion by identifying effective public financial management (PFM) institutions that play a critical role in supporting the implementation of policies for national development and poverty reduction [5]. Furthermore, PEFA emphasizes the importance of comprehensiveness and transparency in budget and fiscal risk supervision, ensuring that fiscal and budget information is accessible to the public, thus ensuring comprehensive, consistent, and accessible information for users [6].

In 2023, Transparency International Brazil [TIBR] released the Transparency and Public Governance Index, which analyzed the Legislative Assemblies of each state and the Federal District in Brazil [7]. The aim of the index was to evaluate the level of transparency and public governance in each state capital. The results of the index were concerning, as only four out of the 27 state capitals and the federal district were rated as having a good level of transparency and public governance. Twelve were rated as fair, eight were bad, and three were very poor [7].

In response to this, the Federal Comptroller General (CGU) presented the Transparent Brazil Scale - 360º Assessment, an innovative methodology that evaluates public transparency by assessing both active and passive transparency [8]. It covers several aspects, including revenue and expenditure, tenders and contracts, administrative structure, public servants, and monitoring of public works. This assessment aims to help monitor the actions implemented by national, subnational, and local governments, and promote access to information. The Transparent Brazil Scale - 360º Assessment is expected to make a significant impact in improving transparency and public governance in Brazil. By analyzing various aspects of public governance, it will help identify areas that need improvement and ensure that government officials are accountable for their actions. Through this assessment, people will have access to reliable information and be able to hold their elected officials accountable [8].

According to Barbosa et al., digital technologies have strongly impacted the relationship between voters and politicians, expanding channels of interaction and relationship [9]. Despite often not reflecting the real public or the basis of political action, this demand results in greater proximity and popular pressure. E-gov has been boosted in recent years, mediating the relationship between citizens and the government, and encouraging debates on the impact of design and usability on the population's dialogue with public management. Usability focuses on how a system is used and can impact user characteristics, especially cognition, the ability to act, and perceive responses [9].

It has been observed that much research in the field of transparency ignores the means through which information is accessed. For instance, in a vast and diverse country like Brazil, it is important to consider the type of device and connection available to citizens. The Regional Center for Studies for the Development of the Information Society (Cetic.br) conducted research in 2023 to measure the possession, use, access, and habits of Brazilians regarding ICT [10]. The study revealed that 62% of Brazilians mainly use their smartphones to access the internet. When the data was stratified, it was found that women (64%), black and mixed-race individuals (63% and 67%, respectively), and those belonging to the DE classes (84%) primarily use the internet via cell phones, highlighting the importance of this device in the country.

Cetic.br also found that 36 million Brazilians do not use the internet. This group is larger among urban residents (29 million), those with an elementary school education or less (29 million), black and brown people (21 million), those in the DE classes (19 million), and individuals aged 60 or older (18 million). The primary reasons cited for not accessing the internet were lack of interest (35%) and lack of computer skills (26%). Interestingly, when considering only class A, the "lack of interest" jumps to 90%, while the "lack of ability" increases when the age group of 35 to 44 years is considered, rising to 45% [10].

In their studies, Jácome Filho & Macêdo discussed the responsiveness of public transparency websites in Paraíba on mobile devices. They concluded that the lack of compatibility with mobile devices hinders the popularization of access to government information [11]. The authors also found that this limitation impedes the population's access to public information, resulting in factors that impede transparency. This is because most of the population accesses such data via cell phones. There are other studies relating to Human-Computer Interaction (HCI) and usability, especially in the Nielsem Heuristics approach, which supports this article.

This text discusses the issue of transparency in Brazil and the means by which citizens can access information. It highlights the fact that a large percentage of the population in Brazil relies on smartphones as their primary means of accessing the internet. However, many Brazilians still do not have access to the internet, particularly those living in urban areas with lower levels of education and belonging to lower socio-economic classes.

The authors of the text cite a study that shows that lack of interest and lack of computer skills are the main reasons why some people do not have internet access. The study also highlights the fact that public transparency websites in Brazil are not always compatible with mobile devices, which can hinder people's access to government information.

To address this issue, the authors of the text conducted an evaluation of public transparency websites in Brazilian capitals and the federal district. They used heuristics

for mobile devices to analyze these websites and determine whether they were user-friendly and in compliance with legal requirements. The study was conducted using a deductive method and had an exploratory nature.

Overall, the text highlights the importance of ensuring that public information is accessible to all citizens, regardless of their socio-economic status or level of education. By making public transparency websites more user-friendly and compatible with mobile devices, governments can promote transparency and improve their relations with society.

2 Heuristics and Interface Evaluation

In the realm of software evaluation, there are a variety of methods available to assess its quality and effectiveness. These methods are designed to evaluate various aspects of software use, including usability, accessibility, and user experience. To achieve this, evaluators are provided with guidance on how to collect and analyze data [9].

The evaluator can choose the appropriate method based on the evaluation requirements and available resources. For instance, if the evaluator needs to understand how users interact with the software in a natural setting, they may prefer a user observation method. On the other hand, if the evaluator needs to assess the software's compliance with specific guidelines, they may prefer an inspection method.

When it comes to evaluating human-computer interaction (HCI), there are three main categories of evaluation methods: investigation, user observation, and inspection methods. Investigation methods involve collecting data through surveys, interviews, and focus groups. User observation methods involve observing users as they interact with the software to determine their behavior and preferences. Inspection methods involve analyzing the software's design and code to identify potential issues that could affect its usability and user experience.

In summary, by utilizing these evaluation methods, software evaluators can gain valuable insights into how users interact with software and identify areas for improvement to optimize the user experience [9].

The inspection method aims to predict certain user experiences and identify design flaws, allowing the evaluator to project possible outcomes of that experience. Observation methods analyze user data about situations carried out in a system, identifying real problems that users face during their usage experience. Research methods include direct contact with users through interviews, field studies, questionnaires, among others, providing the evaluator with access to the user's opinions, expectations, and behavior in relation to the system [9, 12].

Inspection methods allow the evaluator to examine a solution to anticipate the possible consequences of certain design decisions [9, 12]. As it does not involve users, this method aims to identify problems that may arise while using the system. By examining different designs and comparing them, it is possible to find flaws in design and experiences and suggest solutions for product improvements. When examining an interface, evaluators assume the role of a user with a certain profile, with knowledge and experience in certain activities. Their work is to identify problems that the user may encounter when interacting with the system and what support is available to help them overcome these problems.

Heuristic evaluation is a usability inspection technique that suggests the evaluator navigate through the system interface to search for possible problems based on a set of heuristics such as those proposed by Nielsen [13]. This type of evaluation is effective in finding different types of usability flaws. It is also a quick and low-cost evaluation alternative when compared to empirical methods [9].

Heuristic evaluation is a valuable technique for improving the design and functionality of systems, whether they are still in the development phase or already in commercial use. This method can be applied throughout the development process, from prototyping to final software testing before launch. For systems that are already in use, the goal of heuristic evaluation is to enhance the system's performance and user experience through future versions.

Nielsen's heuristics describe ten desirable characteristics of interaction and interface that a system should possess. These heuristics are designed to ensure that the system is user-friendly, intuitive, and meets the needs of the user [9, 13].

The first heuristic (System visibility and status) emphasizes the importance of providing timely feedback to the user about their actions. The second heuristic (Conformity of the system with the real world) focuses on using a language that is familiar to the user and adheres to real-world conventions. The third heuristic (User control and freedom) emphasizes providing users with an easy exit mechanism, allowing them to leave the system without any significant difficulties.

The fourth heuristic (Consistency and standards) stresses the importance of consistency in the system, ensuring that different actions or situations do not mean the same thing and that the same element has the same meaning across different situations. The fifth heuristic (Recognition rather than retrieval) is focused on minimizing the user's memory load by making objects, actions, and options visible.

The sixth heuristic (Flexibility and efficiency of use) emphasizes the importance of optimizing the experience for both experienced and less-experienced users. The seventh heuristic (Aesthetics and minimalist design) emphasizes the importance of using minimalist design and avoiding irrelevant or unnecessary information on the screen.

The eighth heuristic (Error prevention) stresses the importance of preventing errors from occurring. The ninth heuristic (Help users recognize, diagnose, and recover from errors) emphasizes the importance of simple language to expose errors that have occurred and suggest a solution to help users recognize, diagnose, and recover from errors.

The tenth heuristic (Help and documentation) emphasizes the importance of providing users with easy-to-find information and instructions that are focused on the user's task, list concrete steps to be performed and are not too extensive.

Heuristic evaluation is typically conducted by a team of evaluators who explore the system by comparing the design with the heuristics. Once the process is completed, an evaluation report is generated which describes the design errors or violated heuristics with a degree of severity and suggestions for improvements [9, 12–14].

In 2011, the Organization for Economic Cooperation and Development (OECD) conducted a study that showcased the evolution of government solutions. The study highlighted that the development of services is closely linked to the explosion of wireless access points and the expansion of access to smartphones, which has widened the possibility of accessing government services [15]. The study also pointed out that as the

electronic government (e-government) evolved, it paved the way for the advancement of mobile government (m-government). This led to the desire for accessing services through applications rather than just web browsers [15]. In this current context, Machado Neto and Pimentel developed heuristics focused on evaluating mobile devices based on the initial heuristics proposed by Nielsen [13, 14], as shown in Table 1.

Thus, Machado Neto and Pimentel argue that heuristic proposals can, for mobile devices, expand the performance of the concepts proposed by Nielsen, overcoming cosmetic problems, and adapting to new aspects observed from contributions by other authors, as reported in the construction of Table 1 [15]. In this sense, the authors in their results suggest that heuristics for mobile devices are more easily adapted to the evaluation of mobile devices and that because they are the main portals for accessing the internet, they should be part of such an evaluation.

The overall analysis based on Nielsen usability heuristics highlights both significant strengths and areas of improvement among the evaluated sites. Among the positive aspects, it is notable that many websites are well-responsive, adapting to different devices and screen sizes. Furthermore, visual consistency between web and mobile versions is observed in several cases, which contributes to a more cohesive and intuitive user experience, also minimizing user memory overhead. The presence of accessibility options, such as widgets for Brazilian Sign Language (Libras) and text-to-audio transcription, is a positive point, reflecting the commitment to digital inclusion.

However, several problematic points require attention. Contrast issues between text and history have been identified on several websites, which can make reading difficult for users with visual impairments. Furthermore, overlapping elements, especially when opening menus or expanding functionality, impair the clarity and usability of the website. The lack of information in certain areas, both in the web and mobile versions, is a critical point that compromises the user experience and the consistency of information.

The organization of information in the mobile version is also an area of concern. Some websites face difficulties in optimizing screen space, resulting in cluttered interfaces and navigation difficulties. The lack of accessibility options in the mobile version in some cases contrasts with the presence of these options in the web version, which limits accessibility for certain users. The presence of low-resolution images harms the visual quality and can impact the understanding of the content.

In terms of consistency and standardization, some websites have problems when accessing different titles or pages within the domain, which confuses users and impairs the browsing experience. Websites must maintain a consistent and predictable visual identity in all interactions to ensure user familiarity and trust.

In short, heuristic assessments highlight the need for continuous improvements in areas such as contrast, information organization, accessibility, and visual consistency to ensure a more positive and inclusive user experience across platforms. Attention to these details will significantly contribute to improving the usability and effectiveness of the sites evaluated.

2.1 Public Transparency

In Brazil, public transparency is a fundamental aspect of the country's 1988 Federal Constitution. This constitution encourages the public's participation in the development

Table 1. Heuristics for usability evaluation of mobile device interfaces: second version [15].

Heuristic Code	Description
HM1 - Screen space usage	The interface should be designed so that items are neither too far away nor too stuck. Margin spaces may not be large on small screens to improve information visibility. The more related the components are, the closer they should appear on screen. Interfaces should not be overloaded with a large number of items
HM2 - Consistency and standards	The application must keep components in the same place and with the same appearance throughout the interaction, to facilitate learning and stimulate the user's short-term memory. Similar functionality must be performed by similar interactions. The metaphor for each component or feature must be unique throughout the application to avoid misunderstandings
HM3 - Visibility and easy access to all information	All information must be visible and legible, both in portrait and landscape. This also applies to media, which must be fully displayed unless the user chooses to hide it. Elements on the screen must be properly aligned and contrasted
HM4 - Adequacy of the component to its functionality	The user must know exactly what information to insert into a component, without ambiguities or doubts. Resource metaphors should be understood without difficulty
HM5 - Adequacy of the message to the functionality and user	The application must speak the user's language in a natural and non-invasive way so that the user does not feel pressured. Instructions for executing functionalities must be clear and objective
HM6 - Error prevention and fast recovery to the last stable state	The system must be able to predict a situation that leads to a user error based on some activity already carried out by the user [8]. When an error occurs, the application should quickly notify the user and return to the last stable state of the application. In cases where returning to the last stable state is difficult, the system must transfer control to the user, so that he or she can decide what to do or where to go

(*continued*)

Table 1. (*continued*)

Heuristic Code	Description
HM7 - Ease of entry	The way the user provides data may be based on assistive technologies, but the application must always display the input data legibly so that the user has full control of the situation. The user must be able to provide the necessary data in a practical way
HM8 - Ease of access to all features	The application's main features must be easily found by the user, preferably in a single interaction. The most frequently used functionalities can be performed through shortcuts or alternative interactions. No functionality should be difficult to find in the application interface. All input components must be easily assimilated
HM9 - Immediate and observable feedback	Feedback must be easily identified and understood so that the user is aware of the status of the system. Local updates on the screen should be preferred over global ones because the latter maintains the status of the interaction. The interface should give the user the option to hide messages that appear repeatedly. Long tasks must offer the user a way to perform other tasks simultaneously with the task being processed. Feedback should have a good tone, be positive, and not be redundant or obvious
HM10 - Help and documentation	The application must have a help option where common problems and ways to resolve them are specified. The problems considered in this option should be easy to find
HM11 - Reduced user memory load user memory load	The user should not have to remember information from one screen to another to complete a task. The interface information must be clear and sufficient for the user to complete the current task

of public policies and is commonly referred to as the Citizen Constitution. According to the Principle of Publicity outlined in article 5, item XXXIII of the constitution, every citizen has the right to obtain information of interest to them. However, certain information that is essential to the security of society and the State may be kept confidential [17].

To reinforce this legal framework, two complementary laws were enacted: Complementary Law 131/2009 (Transparency Law) and Law no. 12,527/2011 (Access to

Information Law or LAI) [2, 3]. The Transparency Law amended the Fiscal Responsibility Law (LRF) and requires detailed data on the budgetary and financial execution of the Union, the States, the Federal District, and the Municipalities to be made available. The LAI guarantees citizens' right to request information from the public administration, and the rule is that transparency should be the norm, and secrecy the exception.

A study by Machado, Nalini, and Machado analyzed the impact of the LAI on the behavior of public supervisory agents. They found that the law led to a decrease in the number of cases, but the penalties imposed were more severe. Therefore, their research supported the hypothesis that the LAI impacted the behavior of public inspection agents regarding the imposition of sanctions [18].

However, researchers evaluated whether the information made available through public transparency means has quality, utility, and sufficiency. Members of the Social Observatory of Brazil (OSB) believe that the quality of information does not meet the objectives of the Access to Information Law [19]. Although public transparency tools have been useful in reinforcing social control, the information available is not sufficient to exercise social control.

Cruz, Silva, and Santos argue that transparency should be a guiding principle for all activities carried out by public managers [21]. This ensures that citizens have access to and understanding of what government managers have accomplished. The public administration must demonstrate the application of resources to meet the collective interests of society. Transparency in management ensures that the population can see whether the proposed policies are being implemented and allows them to monitor the trust placed in the manager's governability.

It is absolutely imperative for administrators to strictly adhere to legal guidelines and regulations in order to ensure efficient management of local resources [21]. Any deviation from these guidelines can lead to severe negative consequences, such as the blocking of voluntary transfers of federal resources and a significant reduction in the ability to attract external investments. These outcomes can cripple local economic growth and make management processes inefficient to the point of being completely unsustainable.

By meticulously following legal guidelines, administrators can create a stable and secure environment for local businesses to operate and thrive, providing them with the necessary resources and support to grow and succeed. This, in turn, can significantly boost the economic health of the region, creating new opportunities for employment, investment, and community development [22].

In conclusion, there can be no compromise when it comes to upholding legal guidelines. Administrators must remain vigilant and up-to-date on all relevant regulations and laws, ensuring that they are fully compliant with all requirements. Failure to do so can have dire consequences for the entire community, making it imperative that legal guidelines are followed strictly and without fail.

After conducting a thorough analysis of the implementation of public transparency in Brazilian municipalities, it was found that the level of transparency with respect to mandatory information is extremely low [23]. This indicates that there is a significant gap between the existing legal framework and its actual implementation in practice. The results of this study highlight the urgent need for further research to examine whether the existing legal requirements are being met by the municipalities in practice, and to

identify the factors responsible for the low level of transparency. Such research can help to identify the gaps in the existing legal framework and suggest ways to improve the transparency and accountability of the public institutions.

3 Methodology

The research conducted here utilized an experimental strategy based on the inductive method, which is a research method that involves the collection of data in order to generate hypotheses. This approach was chosen because it allowed for the exploration of new ideas and the generation of new insights. In addition, the study employed exploratory features to provide a more comprehensive understanding of the topic under scrutiny.

Furthermore, a survey strategy was implemented for mobile devices, given their high level of representativeness. This enabled the researchers to collect data from a diverse range of participants and to obtain a more accurate understanding of the subject matter.

The decision to use a mobile survey strategy was also based on the participants' willingness to answer a questionnaire at the end of their evaluations. This was done to learn more about the extent of transparency and the impact of vehicles that promote it. Overall, this research was conducted with a high level of rigor and attention to detail in order to provide valuable insights into the topic under investigation.To evaluate the devices and mobile access based on heuristics, we used the theoretical constructs of Machado Neto and Pimentel [16], based on Nielsen [13], Williams [24], Shneiderman and Plaisant [25], Bertini, Gabrielli and Kimani [26], Dix et al. [27], Moraveji and Soesanto [28].

The interface evaluation included mobile devices using a strategy based on heuristics in the 27 local government sites in the Brazilian capitals and the federal district. This sample is representative, as it is made up of the most important city in each state in Brazil. On the other hand, it promotes a vision capable of contemplating those more developed and those less or even developing.

Following the proposed idea, research participants had knowledge of heuristic evaluation, participated in the Interaction and Media Laboratory at the Federal University of Paraíba (UFPB), and worked with the application of this method to transparency sites. Two evaluators were chosen to propose a longitudinal analysis with the sites participating in the research.

The analysis of the results used the responsiveness analysis strategy, covering the following steps: (a) access to the website via a mobile device; (b) search for information vertically; (c) placing the device horizontally and checking responsiveness, following the proposed checks [15]; and (d) classification of sites as responsive, partially responsive, and non-responsive.

This classification was demonstrated in tables, and a comparison of the results in the region and country was also promoted, aiming to offer the reader a descriptive statistical view of the results.

4 Results

This section presents and discusses the results obtained from analyzing electronic sites. An analysis was conducted on the transparency sites of the 27 local governments of the state capitals and the federal district. This was done through a web tool, which resulted in an assessment of the scope of transparency of these sites. Therefore, this survey provides observations that cover the entire investigated population.

Table 2. Heuristic evaluation data.

Capitals	Estate	Region	General Diagnostics of the Site	Frequency in the Region	Frequency in Brazil
Brasília*	DF	Midwest	Responsive with Mobile Version	25,00%	3,70%
Goiânia	GO	Midwest	Partially responsive	25,00%	3,70%
Cuiabá	MT	Midwest	Partially responsive	25,00%	3,70%
Campo Grande	MS	Midwest	Partially responsive	25,00%	3,70%
Maceió	AL	North East	Not responsive, distorts when turning the screen	11,11%	3,70%
Salvador	BA	North East	Not responsive, distorts when turning the screen	11,11%	3,70%
João Pessoa	PB	North East	Responsive	11,11%	3,70%
Fortaleza	CE	North East	Partially responsive	11,11%	3,70%
Recife	PE	North East	Partially responsive	11,11%	3,70%
Teresina	PI	North East	Partially responsive	11,11%	3,70%
Natal	RN	North East	Partially responsive	11,11%	3,70%
Aracaju	SE	North East	Partially responsive	11,11%	3,70%
São Luís	MA	North East	Partially responsive – horizontally	11,11%	3,70%
Rio Branco	AC	North	Not responsive, distorts when turning the screen	14,29%	3,70%
Macapá	AP	North	Not responsive, distorts when turning the screen	14,29%	3,70%
Manaus	AM	North	Not responsive, distorts when turning the screen	14,29%	3,70%

(*continued*)

Table 2. (*continued*)

Capitals	Estate	Region	General Diagnostics of the Site	Frequency in the Region	Frequency in Brazil
Porto Velho	RO	North	Responsive	14,29%	3,70%
Belém	PA	North	Partially responsive	14,29%	3,70%
Boa Vista	RR	North	Partially responsive	14,29%	3,70%
Palmas	TO	North	Partially responsive	14,29%	3,70%
Vitória	ES	Southeast	Partially responsive	25,00%	3,70%
Belo Horizonte	MG	Southeast	Partially responsive	25,00%	3,70%
Rio de Janeiro	RJ	Southeast	Partially responsive	25,00%	3,70%
São Paulo	SP	Southeast	Partially responsive, distorts when turning the screen	25,00%	3,70%
Porto Alegre	RS	South	Not responsive, distorts when turning the screen	33,33%	3,70%
Curitiba	PR	South	Partially responsive	33,33%	3,70%
Florianópolis	SC	South	Partially responsive	33,33%	3,70%

It can be observed from Table 2 that government websites in the capitals and the federal district are only partially responsive, which means they fail to meet the evaluation heuristics criteria as predicted by the theoretical basis. The assessment is carried out by region, and it shows that the central-west region, which includes three capitals and the federal district, has partly responsive websites covering 75% of the region. The federal district's website is responsive and has a mobile version.

The northeast region of Brazil has 22.22% non-responsive websites, 11.11% responsive, and 66.67% partially responsive. Non-responsive websites distort when turned horizontally, and one partially responsive website also presents an error in the horizontal position. Most government entities use this feature to enlarge the screen and facilitate reading.

In the northern region, non-responsive and partially responsive websites reached 42.86%, which limits transparency. Only one website proved to be responsive, representing 14.29% of the region. The southeast region has all partially responsive websites, with no capital being responsive. In the southern region, one website is non-responsive, and two are partially responsive.

The evaluation was also carried out for the entire population, where 22.22% were considered "not responsive, distorts when turning the screen," 11.11% were responsive,

and one had a mobile version which improved its responsiveness. The remaining 66.67% were partially responsive.

Analyzing the capitals, we found that:

(1) Rio Branco (Acre): Showed excellent responsiveness of the elements on the main page. No news and other elements were cut off on this screen. When opening the sandwich menu, to access other features of the site, elements were noticed overlapping texts. The website presents consistency of visual standards between the web and mobile versions. Documentation and informative texts have a font size that allows users to read clearly on both platforms. The main information is transparent to the user in both versions. The use of screen space was well implemented.
(2) Maceió (Alagoas): Excellent responsiveness. Accessibility for people with visual problems, with accessibility widget for Brazilian Sign Language. Excellent consistency in visual standards between the mobile and web versions. Clear and easily accessible information. No problems with button overlap were noticed. Vibrant colors and no contrast problems.
(3) Macapá (Amapá): Consistent layout between web and mobile versions. Serious responsiveness issues, with many buttons and texts being cut off and overlapping. Problems with contrast between text and background color. Some fill fields have lower opacity, contributing to contrast issues. Main information was accessible through the menus. Good documentation and easy access.
(4) Manaus (Amazonas): Good responsiveness. Accessible information. Low contrast in some texts concerning the background. Consistency of standards between web and mobile versions. Good accessibility options, with text-to-audio transcription and Brazilian Sign Language. Good documentation. Good user control options, with easy access to previous screens. No problems with overlapping buttons or cutting visual elements were noticed.
(5) Salvador (Bahia): Easy access to information. Good responsiveness. No contrast problems were noticed. Consistency of standards between web and mobile versions. Large fonts make it easier to read on mobile. Little use of screen space, with icons and other elements that are too large, taking up a lot of space. Accessibility options.
(6) Fortaleza (Ceará): Good responsiveness. Some images were not well-resized for mobile, breaking some standards between web and mobile versions. Low contrast problems in texts and icons can make reading difficult for people with low vision. In the mobile version, there is a lack of information, compared to the mobile version, generating information consistency problems.
(7) Federal District: Presents responsiveness problems. Texts invading other sections. Little information on the main page. The information takes up a lot of screen space, causing screen usage problems. Good contrast. Good error messages.
(8) Vitória (Espírito Santo): Good responsiveness. The mobile version has better organization of information and visual elements than the web. There is a serious problem with the items listed in the sandwich menu, as it is not possible to scroll down without opening an item on the list, making it impossible to reach lower options on the list. It presents a good contrast. There are many icons and little text, making it difficult to understand the redirection when clicking these buttons.
(9) Goiânia (Goiás): Good responsiveness. There is a lack of information compared to the web version. Good contrast. Good use of the screen. Good error messages.

Saves the user's last search if they leave the search field, helping to minimize the user's memory requirements. No problems with overlapping buttons or cutting visual elements were noticed.

(10) São Luís (Maranhão): Excellent responsiveness. Good image resizing. Consistency of standards from the web version to mobile. There is no loss of information when porting to a mobile platform. Clear and findable information. Good error messages. Very well-organized information and visual elements. Accessibility options. Good error messages.
(11) Cuiabá (Mato Grosso): Excellent user feedback, good page loading status. Excellent customization for font size, contrast, accessibility, allowing user control. Good responsiveness. There are some screen usage problems, for example, when opening the sandwich menu, the menu occupies the entire device screen. Good documentation, the icons are accompanied by text. No problems with overlapping buttons or cutting visual elements were noticed.
(12) Campo Grande (Mato Grosso do Sul): Many accessibility options, allowing greater control to the user. Good responsiveness. Icons accompanied by text. Good contrast. No problems with overlapping buttons or cutting visual elements were noticed. Consistency and standards problems concerning the web version, as there is a lack of information.
(13) Belo Horizonte (Minas Gerais): Good responsiveness. Good consistency and standards concerning the web version. Good accessibility options. Images accompanied by text. Good use of the screen. Good system status feedback. No problems with overlapping buttons or cutting visual elements were noticed.
(14) Belém (Pará): Average responsiveness. There are problems with the spacing between buttons. In the mobile version, no accessibility options were found, whereas they exist in the web version. There are contrast issues in the text compared to the background color. Low-resolution images. Bad use of the screen in the sandwich menu. Layout consistency issues when accessing certain pages within the website.
(15) João Pessoa (Paraíba): Good responsiveness. Low-contrast images. Texts in low contrast. There are no accessibility options. Good error messages. Consistency and standards problems depend on the page you access within the domain.
(16) Curitiba (Paraná): There are responsiveness problems. When rotating the screen, there is no good arrangement of information on the screen, taking up a lot of space. Visual effects are wrongly amplified in the mobile version. There are good accessibility options. Good error messages. Good system status feedback. Good documentation, images always accompanied by text, and good font size.

5 Conclusions

The study conducted has revealed that upon analyzing several websites, some of them require improvements or even a complete reconstruction, despite some of them having advanced responsiveness. The study has also concluded that the lack of responsiveness with mobile devices is a significant issue that restricts access to public transparency. It makes it difficult for people to search for more detailed information through desktop-type devices, and it creates barriers for the population that has restrictions on access mechanisms.

The study has shown a significant improvement from the previous study conducted by Jácome Filho & Macêdo [11]. The improved coverage of local governments and state agencies that are closer to the population meets the first needs of the people. As a result, most websites are responsible for generating access and making public transparency possible. This is consistent with the findings of Macêdo, Becker & de Sá [20], who have also found that the lack of compatibility in mobile devices makes it difficult to popularize access. This is particularly challenging for people who have limited access to these devices.

Despite the study's findings, there are some limitations that need to be acknowledged, such as the subjectivity of the heuristic evaluation and the limited sample size. Therefore, it is recommended to expand the study to cities with a population of more than 500,000 to increase the study population. This will help in providing more accurate and reliable findings.

References

1. Brazil. Complementary Law n. 101, of May 4, 2000. Establishes norms of public finances focused on responsibility in fiscal management and other provisions (2000)
2. Brazil. Complementary Law n. 131, of May 27TH, 2009. Adds provisions to the Complementary Law no. 101, of May 4, 2000, which establishes public finance norms aimed at fiscal management responsibility and other provisions, in order to determine the availability, in real time, of detailed information about the budgetary and financial execution of the Union, the States, the Federal District, and the Municipalities (2000)
3. Brazil. Law n. 12.527, of November 18, 2011. Regulates the access to information provided for in item XXXIII of art. 5, in item II of § 3 of art. 37 and in § 2 of art. 216 of the Federal Constitution; amends Law No. 8112 of December 11, 1990; revokes Law No. 111 of May 5, 2005, and provisions of Law No. 8159 of January 8, 1991; and makes other provisions (2011)
4. Dias, L.N.S., Aquino, A.C.B., Silva, P.B., Albuquerque, F.S.: Outsourcing of fiscal transparency portals by municipalities. J. Account. Organ. **14**, e164383 (2020)
5. Public Expenditure and Financial Accountability (PEFA). 2011 Public Finance Management - Performance Assessment Framework (2011). https://www.pefa.org/sites/pefa/files/resources/downloads/PMF%20Portuguese_HGRFinal.pdf. Accessed 20 April 2023
6. Public Expenditure and Financial Accountability (PEFA). 2019 Global Report on Public Financial Management (2019). https://www.pefa.org/sites/pefa/files/resources/downloads/2020002207PORpor002_Main%20text.pdf
7. Transparência Internacional Brasil [TIBR]. Índice de Transparência e Governança Pública (2023). https://indice.transparenciainternacional.org.br/
8. Controladoria-Geral da União (CGU). Avaliações independentes EBT - Avaliação 360° - 2ª Edição (2021). https://mbt.cgu.gov.br/publico/avaliacao/escala_brasil_transparente/66
9. Barbosa, S.D.J., Silva, B.D., Silveira, M.S., Gasparini, I., Darin, T., Barbosa, G.D.J.: Human-Computer Interaction and User Experience. Auto publicação (2021)
10. Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br). Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros – TIC Domicílios 2022
11. Jácome Filho, E., Macêdo, J.M.A.: Analysis of responsiveness and usability in websites serving public transparency in a mobile environment: case study in the state of paraíba through heuristic evaluation. In: Kurosu, M. (eds.) Human-Computer Interaction. User Experience and Behavior. HCII 2022. LNCS, vol.13304. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05412-9_8

12. Preece, J., Sharp, H., Rogers, Y.: Interaction design: beyond human-computer interaction. Wiley (2015)
13. Nielsen, J.: Usability inspection methods. In: Conference Companion on Human Factors in Computing Systems, pp. 413–414 (1994)
14. Nielsen, J.: Severity ratings for usability problems. Papers Essays **54**, 1–2 (1995)
15. Organisation for Economic Co-operation and Development: M-government: mobile technologies for responsive governments and connected societies. OECD Publishing, Paris (2011)
16. Machado Neto, O., Pimentel, M.D.G.: In Proceedings of the 19th Brazilian Symposium on Multimedia and the Web, pp. 93–96 (2013)
17. Brazil. Constitution of the Federative Republic of Brazil. 1988
18. De Souza, L., Machado, L.E., Nalini, G., Machado, M.R.R.: Lei de Acesso à Informação (LAI) e Comportamento de Agentes Fiscalizadores de um Tribunal de Contas. Sociedade, Contabilidade e Gestão **17**(1), 43–66 (2022). https://doi.org/10.21446/scg_ufrj.v0i0.42410
19. Baldissera, J.F., Walter, S.A., Fiirst, C., Asta, D.D.: The perception of social observatories on the quality, utility and sufficiency of public transparency of brazilian municipalities. society, accounting and management SCG journal. Rio de Janeiro **14**(1) (2019) (2019)
20. Macêdo, J.M.A., Becker, V., Feliciano de Sá, F.M.: Heuristic-based evaluation of transparency websites of the municipal governments viewed on web and mobile browsers. In: Kurosu, M., Hashizume, A. (eds.) Human-Computer Interaction. HCII 2023. LNCS, vol 14012. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35599-8_29
21. Cruz, C.F., Silva, L.M., Santos, R.: Fiscal management transparency: a study of the websites of the principal municipalities of Rio de Janeiro State. J. Account. Manage. Govern. **12**(3) (2009)
22. de Melo, F.J.A.: The importance of public transparency laws as a tool for management and social control: a study of the official websites of municipalities in Alagoas (2017)
23. Alves, J.F., Miranda, A.R.A., Teixeira, M. A.C., Souza, P.R.R.D.: Active transparency ranking of municipalities in the Brazilian state of Minas Gerais: evaluation of transparency portals based on the Access to Information Law. Cad. EBAPE.BR **19**(3) (2021)
24. Williams, B.: The non-designers design book: design and typographic principles for the visual novice. Peachpit Press, Non Designer's Design Book (2005)
25. Shneiderman, B., Plaisant, C.: Designing the user interface: strategies for effective human-computer interaction, 5th edn. Addison-Wesley (2009)
26. Bertini, E., Gabrielli, S., Kimani, S.: Appropriating and assessing heuristics for mobile computing. In: Proceedings of the Working Conference on Advanced Visual Interfaces, AVI 2006, pp. 119–126. ACM (2006)
27. Dix, A., Finlay, J.E., Abowd, G.D., Beale, R.: Human-Computer Interaction, 3rd edn. Prentice-Hall, Inc. (2004)
28. Moraveji, N., Soesanto, C.: Towards stress-less user interfaces: 10 design heuristics based on the psychophysiology of stress(chi'12 extended abstracts). Pp. 1643–1648. ACM (2012)

Using EEG and Eye-Tracking to Identify Student Attention in Distance Education

Valdecir Becker, Felipe Melo Feliciano de Sá, Daniel de Queiroz Cavalcanti, João Marcelo Alves Macêdo(✉), Signe Silva, and Paulo Henrique Serrano

Laboratory of Interaction and Media, Informatics Center, Federal University of Paraíba, João Pessoa, Paraíba, Brazil
contato@lim.ci.ufpb.br

Abstract. Distance education has undergone significant development and evolution in recent years. It has become increasingly relevant in comparison to face-to-face teaching. Many technologies have been tested and incorporated into the teaching, learning, and performance assessment. However, one issue that requires greater attention is the mapping of student attention during classes, particularly synchronous ones. This article introduces two human-computer interaction technologies, namely eye-tracking and electroencephalography (EEG), to address this problem. According to user tests, mapping the student's gaze on the computer screen and identifying neural attention patterns can be important tools for planning teachers' pedagogical strategies.

Keywords: Attention · EEG · Eye Tracking · Design Science Research

1 Introduction

Meaningful Learning Theory (MLT) suggests that when we learn something new, it is most effective if we can relate it to our existing knowledge and experiences. This way, we can integrate the new information in a way that is relevant and meaningful to us, rather than just memorizing it superficially without really understanding it. This approach to learning is more beneficial than the superficial approach, where the focus is only on memorization without any real comprehension or connection to existing knowledge [1].

The MLT was developed by educational psychologist David Ausubel. This theory is concerned with the active construction of knowledge by the learner. Some key features of meaningful learning include the connection with prior knowledge - this means that new information is related to what the student already knows and understands. By doing this, the new learning is based on a solid foundation and creates deeper connections. In addition, deep understanding means that the learner not only memorizes information but also comprehends the underlying concepts. [2].

From this perspective, effective learning involves the ability to explain, apply and relate new knowledge in various contexts. The relevance and applicability of new knowledge are crucial in giving it meaning and practicality. This enables learners to apply it in real-world situations, making the learning experience more valuable. Additionally,

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 119–133, 2024.
https://doi.org/10.1007/978-3-031-70439-0_8

meaningful learning involves the cognitive process of reflection, analysis, and critical thinking, which engages the learner in a deeper mental process. Finally, the integration and construction of knowledge means that new information is integrated with previous knowledge, creating a network of interconnected concepts [1].

A classic example of meaningful learning is when someone learns a new scientific concept by relating it to everyday situations or other already known concepts. By making these connections, the learner creates a deeper, more lasting understanding. Educators often seek to promote meaningful learning in their teaching practices by encouraging students to actively participate in the learning process, relate content to their own experiences, and build an authentic understanding of concepts. This contributes to more lasting and meaningful learning over time [2].

An important element of meaningful learning is monitoring student learning. It is of fundamental importance for the teacher to know that students are paying attention in class and not wandering or absent. This monitoring, which in face-to-face teaching can be done with a simple look at the class, in distance education becomes an additional challenge to teaching practice. This problem is even more pressing in synchronous remote classes, which have become common in recent years.

In other words, understanding and mapping the effective results of the teaching and learning process in remote and distance learning, ensuring the addition of new knowledge to previous ones and, consequently, meaningful learning, are a categorical challenge today and give rise to the main question of this study: how can we improve learning by using new resources to map student attention and engagement, enabling greater effectiveness and meaningful distance learning? The central hypothesis of the research is that with the use of human computer interaction (HCI) technologies it is possible to monitor student attention and participation more accurately during remote or distance classes.

This article explores the application of two popular technologies in the HCI field for monitoring students during remote learning. These technologies are Electroencephalography (EEG) and eye-tracking. EEG measures the student's level of attention during classes, while eye-tracking tracks the student's gaze and records points of focus on the computer screen or any shift of attention away from the screen.

Initial tests were conducted with five students, using two different videos. The findings suggest that students have a preference for paying more attention to one video profile and feeling more bored while watching the other. By analyzing this data, educators can customize their teaching methods, adjusting them to individual students' needs, which can lead to more meaningful and engaging learning experiences.

2 Theoretical Reference

Distraction, also known as mental wandering, can be a hindrance in the field of education and the teaching and learning processes. In this study, we define mind wandering as a state where the mind wanders or strays away from the task or focus at hand. It occurs when a person's attention is diverted to thoughts, ideas, or concerns that are not directly related to their current activity or train of thought [3].

Mind wandering can occur in various situations, such as when an individual tries to concentrate on work, studying, or a conversation but instead ends up lost in unrelated

thoughts. This can happen due to boredom, external distractions, mental fatigue, or the intrinsically active nature of the human mind, which frequently jumps between random thoughts.

In education, students' wandering can pose a challenge in the teaching-learning process [4–6]. When teaching in-person, a teacher has various methods for observing a student's level of engagement and focus in learning activities. However, in remote or distance learning situations, interactions are virtual, and new strategies are necessary to assess the effectiveness of student learning. Considering the principles of meaningful learning, two techniques that can help teachers are eye tracking and electroencephalography (EEG).

Eye-tracking technology is a well-known and widely used method in Human-Computer Interaction (HCI) and various other fields, including marketing, cognitive psychology experiments, and entertainment, such as in games [7, 8]. It works by detecting the user's eye movements during interaction, which can be achieved through different tools like webcams, infrared cameras, or head-mounted displays. The data collected through this technology can provide valuable insights about users, such as their attention to a specific object on the screen.

When mapping a user's gaze, the specific points that they look at are called "gaze points" or "specific points in the image". A group of gaze points that are fixed on a certain area for a period of time is called a "fixation" [9]. This region is crucial for our study, as it is during this time that the user's main cognitive processes such as understanding, and memory occur.

The quick movement made between the "fixation" points is called a "saccade". During this movement, which lasts between 30 to 80 ms, visual information is blocked. A group of "fixations" can be grouped together based on proximity, forming a "gaze", which can then be divided into "areas of interest" (AOI) [10]. The amount of time spent on each AOI (known as "dwell time") can determine the level of interest the user has in a particular stimulus on the screen. A longer dwell time can indicate a higher level of interest. One way to observe this phenomenon is through the use of heatmaps, which display on the screen the points that captured the user's attention [8].

The article presents a second technique called EEG, which involves analyzing the spontaneous electrical activity of the brain. The technique is used to detect patterns and abnormalities in brain waves [11]. EEG involves placing electrodes on specific points on the user's skull according to the internationally recognized 10/20 system. This enables the recording of brain waves in the cortex. While the technique has good temporal resolution, meaning it captures information in real time, it has low spatial resolution. It's important to select the relevant information and remove noise when processing the data extracted from the brain using EEG.

Studies have shown that midline-frontal beta waves correspond to individual preferences [12]. This means that the higher the amplitude of oscillation in the beta wave frequency observed through EEG when watching a film trailer, the higher the score given by research participants to films related to the same theme [12]. Additionally, the desynchronization of alpha waves in the left-frontal side of the brain is positively related to the level of pleasure and satisfaction perceived when watching commercials [11]. Finally,

an increase in the power of theta waves in the frontal midline is associated with feelings of pleasure [13].

Using eye tracking and EEG, it is possible to diagnose a user's perception of visual stimuli. This helps us to better understand their preferences and the emotional and cognitive processes that occur naturally as they consume content on a screen. Initially, we can be highly certain about where the student is focusing their gaze, which should be on the computer screen. At the same time, we can measure their attention, engagement with the class content, and mental memorization processes with the EEG.

In other words, EEG and eye tracking can help measure students' engagement and attention. These technologies are aligned with MLT and aim to connect new content with prior knowledge. The use of these resources could optimize the learning process, creating a more adaptive and student-oriented educational environment.

3 Related Work

There have been several research studies that have focused on using the techniques of EEG and eye tracking [14–18]. For instance, GuruTutor [15] is an avatar developed to provide virtual instructions to students. It was tested and developed by analyzing visual focus through eye-tracking and patterns of engagement and comprehension through EEG. Another relevant research study that combined the two techniques was conducted using the Neuroscan SynAmps2 system and Curry 8.0 recording software [18]. In this study, EEG data was collected while a conductive gel was used to reduce the impedance at each electrode. Simultaneously, eye movement data was collected using an infrared video-based eye tracker, the EyeLink Portable Duo from SR Research.

There are several studies that have used eye-tracking technology to identify students' concentration and attention levels during learning activities [16, 18–25]. Some of these studies have used WebGazer, which can detect cognitive states during online reading comprehension tasks. These cognitive states can be related to both task-related and non-task-related comprehension. Other studies have used eye-tracking glasses, which have shown that multimedia technologies need to be carefully chosen for effective learning [20].

Although there are other indicators, such as physiology and gestures, eye-tracking is considered the best short-term indicator for identifying the direction of visual attention. However, detecting mind wandering can be difficult, as it involves "looking without seeing" and directing attention to other content or subjects. Despite this limitation, gaze tracking can effectively identify when the mind begins to wander [24]. This is because conventional eye movement patterns change when someone is mentally distracted. For example, when someone is not focused, they are less likely to hold their gaze on parts of the text they have already read. Additionally, blink frequency increases when the mind is wandering, perhaps because there is less processing of visual information during blinks.

Research in the field of education has explored the use of EEG in addition to eye tracking, using different technologies [18, 22, 26–34]. By using machine learning methods [28, 33, 34], EEG measurements were utilized to predict mind wandering. It was observed that the Support Vector Machine (SVM) model is more effective than the logistic regression model in classifying mind wandering states within and between subjects.

This finding is consistent with previous studies that demonstrate the effectiveness of nonlinear models in determining the boundaries between attentional states.

Real-world environments have also been studied, where recordings of EEG activity during live lectures were analyzed, and power measurements in the theta, alpha and beta frequency bands achieved an average detection accuracy of 80% to 83%. It was discovered that students who demonstrated satisfaction [26, 30, 31] showed greater power in alpha and beta frequencies when subjected to tests in which the video instructor used pointing gestures. These results suggest less sensorimotor involvement in the processing of this information.

Although the topic of mapping attention in the teaching and learning process, with a focus on distance education, is present in scientific research in different parts of the world, no research was found that takes into account both the direction of the gaze and the attention and concentration of students. This data becomes even more relevant if we include the use of a webcam for eye tracking as a requirement.

4 Methods

This article utilized eye-tracking software to identify the user's points of interest in remote classes, while the Emotiv Insight 2.0 headset was utilized to map neural patterns. The research is based on Design Science Research (DSR), which justifies the development of problem-solving-oriented artifacts as a way of generating scientific and technological knowledge [35–37]. DSR was considered suitable for this study after identifying gaps in technical and scientific production involving EEG and eye tracking for attention mapping in remote classes. The methodological process consists of six activities:

1. Identification and motivation of problems: This step aims to define the research problem and justify the value of a solution.
2. Defining the objectives of a solution: The objectives of a solution are inferred, considering the definition of the problem, and analyzing what is possible and viable within the scope of the research.
3. Design and development: At this stage, the artifact is created.
4. Demonstration: The researcher presents the created artifact and shows how it solves the problem, or part of it.
5. Evaluation: In this activity, the researcher observes the use of the artifact and evaluates how well it solves the problem. A comparison is made between the objectives and the results observed during the use of the artifact in the demonstration.
6. Communication: this is the final stage that closes the DSR cycle. In this activity, the problem, along with its context, relevance, and the created artifact's usefulness, quality, and novelty, rigor of the design, and effectiveness must be disclosed to impacted professionals, researchers, scientists, and other relevant audiences.

5 Objectives and Development of the Artifact

Considering Activity 2 of the DSR and the central research problem of identifying attention in distance synchronous classes, the following objectives were established for the artifact:

1. Use of non-invasive, easily accessible, and low-cost devices.
2. Reading and identification of neural wave patterns.
3. Availability and access to real-time data processing and subsequent treatment using proprietary software.
4. Storage of data for future retrieval.
5. Data retrieval for study and analysis of neural patterns.
6. Graphical visualization of neural patterns.
7. Integration of the system with the individuals' focus of attention.
8. Integrated analysis of attention during classes.

An initial description of the system requirements is in [38]. The subsequent sections describe the design, implementation, and testing of the artifact, along with the lessons learned so far.

For this stage, the Emotiv Insight 2.0 headset was used, which meets objectives 1 and 2 entirely and objective 3 partially. A Python script was developed for data processing and visualization, which completes objective 3. The Cortex API [39], developed by the headset manufacturer and available online for free, was also used.

The Emotiv Insight 2.0 [40] is a device that measures EEG signals using wireless Bluetooth technology. It has 5 reading channels located at positions AF3, AF4, T7, T8 and Pz of the international 10/20 system. It also has Driven Right Leg (DRL) connections and Common Mode Sense (CMS) on the left mastoid, which act as the system reference. The device collects samples sequentially at a frequency of 2048 Hz, which are then filtered and reduced to a sampling rate of 128 samples per second. The collected data is processed by a 14-bit analog-to-digital converter, with 2 bits of instrumental noise floor being discarded. One Least Significant Bit (LSB) represents 51μV, and the dynamic range for the inputs is 8400μV. The headset is equipped with a fifth order Sinc filter, digital notch filters for 50 Hz and 60 Hz, AC coupling, and can recognize bandwidths for brain waves in the range of 0.5–43 Hz. The Emotiv Insight is also capable of detecting movements using the accelerometer, magnetometer, and gyroscope present in the ICM-1 IMU attached to the headset.

The user is exposed to different classes using the Emotiv Insight device. The system architecture is shown in Fig. 1. While the electromagnetic pulses are read via the headset, the data is interpreted in a cloud server environment through constant exchange of information with the Cortex API. The connection is made using the Websockets Secure Protocol on a Websocket server and the JSON-RPC 2.0 protocol [41]. After the application authentication process using credentials validated by the Emotiv Launcher application, it is possible to extract different data streams from the readings. The records are immediately displayed in a graph dynamically built in a Python language program using the Matplotlib library. This research presents information using power readings in μV2/Hz in the Alpha, Low Beta, High Beta, Gamma, and Theta frequency bands, transmitted at a frequency of 8Hz, in accordance with the specifications of the aforementioned API [39].

Simultaneously a CSV file is created and fed with the obtained information. The application can be fed back with the generated files to study and analyze the recorded neural patterns. To facilitate more in-depth readings of the extracted data, the graph exploration features of the Matplotlib library are utilized in a similar way to real-time visualization.

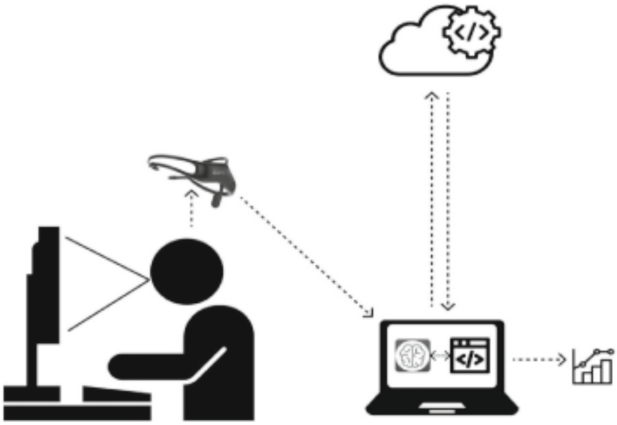


Fig. 1. Diagram of the functioning of the developed artifact [Authoral].

Additionally, the GazeRecorder tool [42], a free online software for eye tracking experiments based on the computer's camera, is used to mitigate the issue of gaze focus and attention. In previous tests, it was observed that emotions vary depending on people and moments, but patterns or causal relationships could not be identified [38, 43]. The relationship between macro and micro expressions can be influenced by the environment, focus, attention, and concentration of the user, resulting in different emotional frames [43]. By allowing synchronization between the readings of brain waves and the direction of the user's gaze, an eye tracking system can evaluate which elements caught the user's attention the most. This information can then be associated with an analysis carried out using the EEG to make more detailed conclusions about what exactly influences the emotion felt during enjoyment.

An experiment was conducted to demonstrate and validate the effectiveness of two video lessons. Five participants, three men and two women aged between 21 and 54, attended the classes which were held in the Laboratory of Interaction and Media. The laboratory was set up to simulate an office environment with a table and a Macbook computer. Prior to the commencement of the classes, participants were asked to fill out pre-test questionnaires to identify their physical and emotional state. Post-test questionnaires were administered at the end of the day to evaluate the two classes attended.

The experiment began with an initial calibration step where participants had to follow a red dot that moved across the screen with their vision. After calibration, the classes were displayed on the screen, and the tool mapped the vision of the participants. This allowed the tool to identify the regions of the video that the participant had the greatest concentration on. Finally, the software generated a heat map over the film scenes, providing the results of the experiment for later analysis. The software used had an accuracy of 1.05°, a precision of 0.129°, and a sampling frequency of 30 Hz [44].

Two videos were used in the experiment, each with a different approach. The first video was a didactic video about how the Solar System moves around the Milky Way. It was divided horizontally into two parts, with the upper part being a 3D animation representing the movement described by the author of the video. The lower part describes

the movements of the animation. The second video was a traditional class without any visual appeal.

Video 1:

Description: This is a didactic video explaining how the Solar System moves around the Milky Way. The video is divided horizontally into two sections. The upper section is a 3D animation representing the movement described by the author, and the lower section describes the movements of the animation.

Author: The video was published by the YouTube channel "aindanaosei" on August 23, 2022.

Duration: The video lasts for 1 min.

Video 2:

Description: This is a didactic video about polynomial equations. The video uses drawings to demonstrate how algebraic calculations are performed with more than one type of variable. The exercise is solved on a sheet of paper.

Author: The author of the video is unknown.

Publication: The publication date and source of the video are unknown.

Duration: The video lasts for 1 min.

6 Outcomes and Discussions

The study aimed to prove that the quality of content produced can be identified or mapped based on brain waves. Five users were tested, with a focus on alpha and beta waves. The alpha waves are closely related to spatial, semantic, and social attention [13, 45–47]. They have several functional correlations that reflect sensory, motor, and memory functions. During physical and mental relaxation with eyes closed, the power levels of this frequency increase. However, the level of alpha waves is reduced during mental or bodily activity with the eyes open. Alpha suppression is a reliable indicator of mental engagement and activity states, particularly during focused attention on various stimuli, which is the focal point of this study. This means that during attentive moments with eyes open, one would expect a low or non-existent rate of alpha waves.

On the other hand, beta waves are related to active, busy, or anxious thinking and high levels of concentration. The level of beta waves becomes stronger as we plan or execute movements, particularly when reaching or grasping requires precise finger movements and focused attention. This increase in beta waves is also noticeable when we observe the body movements of others. Our brain apparently imitates other people's limb movements, indicating that there is an intricate "mirror neuron system" in our brain that is coordinated by beta frequencies [13, 48, 49].

Beta waves can be divided into two categories: low beta and high beta. Low beta waves are associated with negative feelings such as anxiety, fright, and surprise, while high beta waves are linked to positive feelings of surprise, anxiety, and excitement [50]. Therefore, a high rate of high beta waves is expected when the student is enjoying the content and is satisfied with the class.

In the YouTube Shorts video experiment, the first video demonstrated that participants had a greater focus on the 3D animation when it was mentioned in the video author's

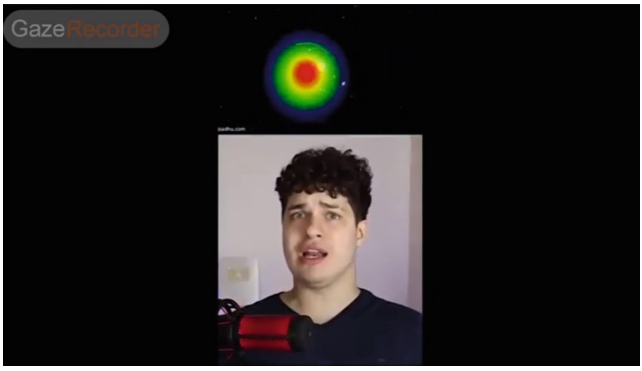


Fig. 2. Eye tracking heat concentration in 3D simulation [Authoral].

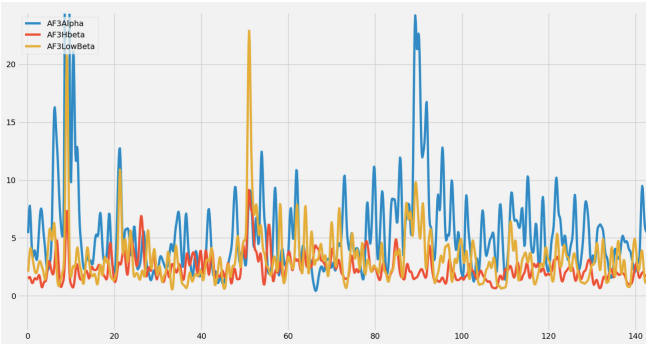


Fig. 3. Low Beta, High Beta and Alpha levels in the left front channel of a specific participant during video 1 enjoyment [Authoral].

speech, resulting in a quick unconscious response from the student. This caused the participants to pay more attention to the content being presented (Fig. 2). When the 3D modeling was not referenced in the video, participants often returned their gaze to the author's face, which detracted from the fact that they were actually paying attention to the content. These results demonstrate that using supporting figures and videos can be an excellent solution to capture student attention.

During the analysis of the brain waves of a user in a video about the movement of the Solar System in the Milky Way, high levels of alpha waves were detected (the blue ones in Fig. 3). The alpha wave can give information about the participant's level of attention. In this case, it can be concluded that the participant was mentally scattered during the class, despite focusing on the graphic elements on the screen. This pattern was observed in all other study participants, and post-test interviews confirmed that this video was preferred due to its interesting content and high-quality production. High beta waves and little gaze wandering were also observed in the participants during the video.

During the video lesson on polynomial equations, researchers observed the participants' eye movements and found that they focused on the specific element that the

teacher referred to. Figure 4 shows that sometimes, the participants' gaze shifted, possibly indicating that they were trying to anticipate the result of the equation. Since the equations were simple, it's likely that the participants became bored or anxious and tried to predict the result with their eyes.

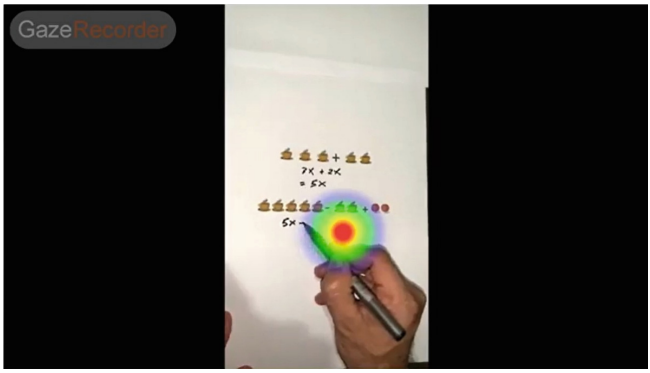


Fig. 4. Eye tracking heat concentration in anticipating results [Authoral].

The video on polynomial equations showed higher levels of low beta waves (in yellow) as seen in Fig. 5. Low beta waves are associated with negative emotions such as surprise, fright, and anxiety, which suggests that users negatively experienced anxiety while watching the video. This could be due to the fact that users already knew what the expected result of the polynomial equation would be, and the video was a slower teaching method without engaging visuals.

In contrast to the first video, the alpha levels in this one were slightly lower, implying that the student might have focused more on an abstract topic. Mental calculations demand greater effort and concentration compared to the mental retrieval of images from graphic elements. The post-test interviews confirmed this fact and indicated that the second video was perceived as less interesting and had less educational value compared to the first one.

Studies have shown that educators can use EEG and eye tracking together to adjust their teaching strategies and materials based on the identified patterns. This approach provides useful data for both synchronous classes and teaching material preparation. For instance, if eye tracking reveals that students tend to focus on certain parts of the text or image, while EEG indicates high levels of brain activity associated with attention, educators can emphasize those sections during explanation or discussion in the classroom.

Conversely, EEG signals and eye-tracking patterns can indicate problematic pedagogical strategies if they show moments of distraction, disinterest or mind wandering. In such cases, educators can intervene by incorporating interactive activities or breaks for discussions to engage the students. This data-driven approach promotes more meaningful learning by adapting the content to the student's individual needs and preferences, resulting in deeper and longer-lasting absorption of knowledge.

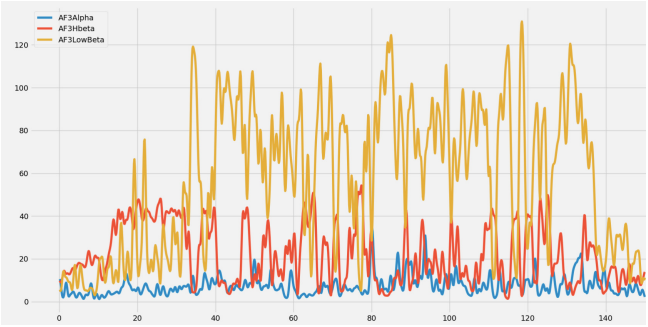


Fig. 5. Low Beta, High Beta and Alpha levels in the left front channel of a specific participant during video 2 fruition [Authoral].

The use of computational processes aligns with the principle of selective attention, which is a fundamental part of the MLT [1]. This theory emphasizes that learning occurs when students focus on relevant and meaningful information. Detailed analysis of fixations and eye movements can help identify which parts of the material are receiving the most attention, enabling educators to adapt teaching methods to emphasize these areas of focus. This, in turn, increases the likelihood of students making connections and developing deeper knowledge foundations. It could also avoid the mind wandering phenomenon.

According to MLT, activating students' cognitive processes requires reflection, analysis, and critical thinking. The presence of these elements can be identified by alpha and high beta waves, which allow for an inference about the student's engagement with the content being taught.

Eye tracking tools have certain limitations when used for experiments as they rely on webcam recordings, which may not be as accurate as recordings made using an infrared camera. Infrared cameras are more suitable for environments where the user is not stationary in front of the screen. Moreover, the free version of the GazeRecorder software has restrictions on the number of tests that can be carried out.

For future work, we recommend obtaining an infrared camera to enhance accuracy in eye-tracking results and integrating open software to analyze it. We also suggest implementing software that can visualize EEG waves. In terms of wave analysis, we suggest the need for more powerful EEG readers that are not prone to interference from hair fibers, like the Emotiv Insight. Additionally, we require more data from a larger number of users to further understand issues related to taste patterns and affinity with the content of the classes as a whole.

7 Final Considerations

The article discussed the use of eye tracking and EEG to monitor students' attention during distance learning. The GazeRecorder website was utilized for eye tracking, while the Emotiv Insight headset and Python software were used to read brain waves. The study aimed to enhance synchronous classes, but the gathered data can also be used to evaluate and redesign teaching materials available asynchronously.

The study confirmed the hypothesis that the use of Human-Computer Interaction technologies, specifically eye-tracking and EEG-based brain-machine interfaces, can accurately monitor student attention and participation during remote or distance classes. The test results showed that neural patterns and eye tracking were compatible with the dynamics of the videos used in the tests. The patterns identified were confirmed by post-test interviews.

It's important to note that the study was conducted in a laboratory and utilized specific technologies, which may not be feasible on a larger scale. While eye tracking based on computer webcams can be replicated on a large scale, using the Emotiv headset may not be practical for larger groups of students. However, one alternative could be the MN8 headset, developed and marketed by Emotiv, which has two reading channels and is available for purchase at $399.00 on the company's website.

Although the study produced results, certain limitations were identified during the research. However, these limitations were mitigated throughout the process. As explained in the text, the visual analysis of the results of waves obtained through electroencephalography remains one of the key challenges faced. It is a manual process that is time-consuming and open to interpretation, which can vary depending on the evaluator.

The Emotiv Insight device has certain limitations that affected the development of the experiment. Although the device is portable and provides a stable connection via Bluetooth, it fails to obtain stable readings in individuals with voluminous or thick fiber hair. This creates difficulties in stabilizing brainwave reading flows. During tests with unstable signal acquisition, an attempt to normalize the results obtained was noticed, either by the device or by the API provided by the manufacturer. This resulted in readings that were incompatible with the participant's actual mental state. Additionally, the type of results obtained, and the format of the output files are limited by the manufacturer's business model. Certain features are blocked unless one subscribes to paid plans, such as obtaining unprocessed EEG data and exporting the results in one's own formats.

A significant limitation is the issue with audio, as students can look at things outside the screen without losing their focus in class. This means that changes in neural patterns may not be identified. Attention remains constant and focused on the class content. However, if something external grabs the student's attention, dopamine levels naturally increase. This shift in attention can be detected through alpha waves, which indicate changes in dopamine levels and their relationship with shifts in focus [51–55]. In the study discussed in this article, no changes in alpha waves were found. One possible reason for this is the short duration of the videos and the characteristics of the tests, which did not cause shifts in attention. This should be taken into account in future tests.

Moreover, it's important to note that the results described are from an initial study intended to highlight potential uses of these technologies in the context of distance learning. The results indicate the potential for mapping student attention and engagement during classes. However, it's important to recognize that there is still work to be done before this research becomes accessible technology for teachers. It's crucial to find ways to visualize the data in an easy and understandable manner for any teacher through an online graphical interface, without having to deal with technical elements related to EEG or eye tracking. This aspect is also a suggestion for future work.

References

1. Moreira, M.A.: A teoria da aprendizagem significativa e sua implementação em sala de aula. Editora Universidade de Brasília, Brasília (2006)
2. Valadares, J.A., Moreira, M.A.: A Teoria da Aprendizagem Significativa: sua Fundamentação e Implementação Almedina (2009)
3. Dias da Silva, M.R., Postma, M.: Wandering minds, wandering mice: computer mouse tracking as a method to detect mind wandering. Comput. Human Behav. **112** (2020). https://doi.org/10.1016/j.chb.2020.106453
4. Wong, J.T., Mesghina, A., Chen, E., et al.: Zooming in or zoning out: examining undergraduate learning experiences with zoom and the role of mind-wandering. Comput. Educ. Open **4** (2023). https://doi.org/10.1016/j.caeo.2022.100118
5. Rebecchi, K., Hagège H.: Educating through attentional states of consciousness, an effective way to develop creative potential? Front. Educ. (Lausanne) **7** (2022)
6. Szpunar, K.K., Moulton, S.T., Schacter, D.L.: Mind wandering and education: From the classroom to online learning. Front. Psychol. **4** (2013)
7. Carter, B.T., Luke, S.G.: Best practices in eye tracking research. Int. J. Psychophysiol. **155** (2020). https://doi.org/10.1016/j.ijpsycho.2020.05.010
8. Duchowski, A.T.: Eye tracking methodology: theory and practice, 3rd edn. Springer International Publishing AG (2017)
9. Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols. In: Proceedings of the Eye Tracking Research and Applications Symposium (2000)
10. Rim, N.W., Choe, K.W., Scrivner, C., Berman, M.G.: Introducing point-of-interest as an alternative to area-of-interest for fixation duration analysis. PLoS ONE **16** (2021). https://doi.org/10.1371/journal.pone.0250170
11. Vecchiato, G., Toppi, J., Astolfi, L., et al.: Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements. Med. Biol. Eng. Comput. **49** (2011). https://doi.org/10.1007/s11517-011-0747-x
12. Lee, Y.Y., Hsieh, S.: Classifying different emotional states by means of eegbased functional connectivity patterns. PLoS ONE **9** (2014). https://doi.org/10.1371/journal.pone.0095415
13. iMotions-Unpack Human Behavior Electroencephalography The Complete Pocket Guide (2019)
14. Huang, Y.M., Liu, M.C., Lai, C.H., Liu, C.J.: Using humorous images to lighten the learning experience through questioning in class. Br. J. Edu. Technol. **48** (2017). https://doi.org/10.1111/bjet.12459
15. Hutt, S., Mills, C., Bosch, N., et al.: Out of the Fr-Eyeing Pan (2017)
16. Hutt, S., Wong, A., Papoutsaki, A., et al.: Webcam-based eye tracking to detect mind wandering and comprehension errors. Behav. Res. Methods (2023). https://doi.org/10.3758/s13428-022-02040-x
17. Atiq, Z., Loui, M.C.: A qualitative study of emotions experienced by first-year engineering students during programming tasks. ACM Trans. Comput. Educ. **22** (2022). https://doi.org/10.1145/3507696
18. Pei, X., Xu, G., Zhou, Y., et al.: A simultaneous electroencephalography and eye-tracking dataset in elite athletes during alertness and concentration tasks. Sci. Data **9** (2022). https://doi.org/10.1038/s41597-022-01575-0
19. Hu, B., Li, X., Sun, S., Ratcliffe, M.: Attention recognition in EEG-based affective learning research using CFS+KNN algorithm. IEEE/ACM Trans. Comput. Biol. Bioinform. **15** (2018). https://doi.org/10.1109/TCBB.2016.2616395
20. Çeken, B., Taşkın, N.: Multimedia learning principles in different learning environments: a systematic review. Smart Learn. Environ. **9** (2022)

21. Hutt, S., Krasich, K., Mills, C., et al.: Automated gaze-based mind wandering detection during computerized learning in classrooms. User Model User-adapt Interact **29** (2019). https://doi.org/10.1007/s11257-019-09228-5
22. Dong, H.W., Mills, C., Knight, R.T., Kam, J.W.Y.: Detection of mind wandering using EEG: within and across individuals. PLoS ONE **16** (2021). https://doi.org/10.1371/journal.pone.0251490
23. Faber, M., Bixler, R., D'Mello, S.K.: An automated behavioral measure of mind wandering during computerized reading. Behav. Res. Methods **50** (2018). https://doi.org/10.3758/s13428-017-0857-y
24. D'Mello, S.K.: Giving eyesight to the blind: towards attention-aware AIED. Int. J. Artif. Intell. Educ. **26** (2016). https://doi.org/10.1007/s40593-016-0104-1
25. Bixler, R., D'Mello, S.: Automatic gaze-based user-independent detection of mind wandering during computerized reading. User Model User-adapt Interact **26** (2016). https://doi.org/10.1007/s11257-015-9167-1
26. Liu, Z., Yu, P., Liu, J., et al.: How do students' self-regulation skills affect learning satisfaction and continuous intention within desktop-based virtual reality? A structural equation modelling approach. Br. J. Edu. Technol. **54** (2023). https://doi.org/10.1111/bjet.13278
27. Xu, X., Sui, L.: EEG cortical activities and networks altered by watching 2D/3D virtual reality videos. J. Psychophysiol. **36** (2022). https://doi.org/10.1027/0269-8803/a000278
28. Dhindsa, K., Acai, A., Wagner, N., et al.: Individualized pattern recognition for detecting mind wandering from EEG during live lectures. PLoS ONE **14** (2019). https://doi.org/10.1371/journal.pone.0222276
29. Chen, C.M., Wang, J.Y., Yu, C.M.: Assessing the attention levels of students by using a novel attention aware system based on brainwave signals. Br. J. Edu. Technol. **48** (2017). https://doi.org/10.1111/bjet.12359
30. Pi, Z., Zhang, Y., Zhou, W., et al.: Learning by explaining to oneself and a peer enhances learners' theta and alpha oscillations while watching video lectures. Br. J. Edu. Technol. **52** (2021). https://doi.org/10.1111/bjet.13048
31. Pi, Z., Zhang, Y., Yu, Q., et al.: Neural oscillations and learning performance vary with an instructor's gestures and visual materials in video lectures. Br. J. Edu. Technol. **53** (2022). https://doi.org/10.1111/bjet.13154
32. Digitalcommons@usu, D., Snowden, A.W., Warren, C.M., et al.: Neural Efficiency and Spatial Task Difficulty: A Road Forward to Neural Efficiency and Spatial Task Difficulty: A Road Forward to Mapping Students' Neural Engagement in Spatial Cognition Mapping Students' Neural Engagement in Spatial Cognition (2021)
33. Jin, C.Y., Borst, J.P., van Vugt, M.K.: Predicting task-general mind-wandering with EEG. Cogn. Affect. Behav. Neurosci. **19** (2019). https://doi.org/10.3758/s13415-019-00707-1
34. Kawashima, I., Kumano, H.: Prediction of mind-wandering with electroencephalogram and non-linear regression modeling. Front. Hum. Neurosci. **11** (2017). https://doi.org/10.3389/fnhum.2017.00365
35. Dresch, A., Lacerda, D.P., Antunes, J.A.V.J.: Design Science Research: Método de Pesquisa para Avanço da Ciência e Tecnologia. Bookman, Porto Alegre (2015)
36. Järvinen, P.: Action research is similar to design science. Qual. Quant. **41** (2007). https://doi.org/10.1007/s11135-005-5427-1
37. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**, 75–105 (2004). https://doi.org/10.2307/25148625
38. Becker, V., Silva, T., Cavalcanti, M., et al.: Potencial das interfaces cérebro máquina para a recomendação de conteúdos em sistemas de vídeo sob demanda. In: 4o Congresso Internacional Media Ecology and Image Studies - Reflexões sobre o ecossistema midiático pós pandemia. Ria Editorial, pp. 145–167 (2021)

39. Getting Started - Cortex API. https://emotiv.gitbook.io/cortex-api/. Accessed 31 Aug 2023
40. Technical Specifications - INSIGHT Manual. https://emotiv.gitbook.io/insight-manual/introduction/technical-specifications. Accessed 31 Aug 2023
41. Connecting to the Cortex API - Cortex API. https://emotiv.gitbook.io/cortex-api/connecting-to-the-cortex-api. Accessed 31 Aug 2023
42. Online Eye Tracking | Webcam eye-tracking software. https://gazerecorder.com/. Accessed 31 Aug 2023
43. Becker, V., Cavalcanti, M., Silva, T., et al.: A System for Graphical Visualization of Brainwaves to Analyse Media Content Consumption. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2022)
44. The comparison of accuracy and precision of eye tracking: GazeFlow vs. SMI RED 250 (2013)
45. Başar, E.: A review of alpha activity in integrative brain function: fundamental physiology, sensory coding, cognition and pathology. Int. J. Psychophysiol. **86** (2012)
46. Rana, K.D., Vaina, L.M.: Functional roles of 10 Hz alpha-band power modulating engagement and disengagement of cortical networks in a complex visual motion task. PLoS ONE **9**, e107715 (2014). https://doi.org/10.1371/JOURNAL.PONE.0107715
47. Başar, E., Schürmann, M., Başar-Eroglu, C., Karakaş, S.: Alpha oscillations in brain functioning: an integrative theory. Int. J. Psychophysiol. **26** (1997)
48. Lainscsek, C., Hernandez, M.E., Weyhenmeyer, J., et al.: Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson's disease from healthy individuals. Front. Neurol. **4** (2013). https://doi.org/10.3389/FNEUR.2013.00200
49. Zhang, J.J.Q., Fong, K.N.K., Welage, N., Liu, K.P.Y.: The activation of the mirror neuron system during action observation and action execution with mirror visual feedback in stroke: a systematic review (2018). https://doi.org/10.1155/2018/2321045
50. Grafton, S.T., Tipper, C.M.: Decoding intention: a neuroergonomic perspective. Neuroimage **59** (2012)
51. Stahl, D., Ferger, B., Kuschinsky, K.: Sensitization to d-amphetamine after its repeated administration: evidence in EEG and behaviour. Naunyn Schmiedebergs Arch. Pharmacol. **356** (1997). https://doi.org/10.1007/PL00005059
52. Ferger, B., Stahl, D., Kuschinsky, K.: Effects of cocaine on the EEG power spectrum of rats are significantly altered after its repeated administration: do they reflect sensitization phenomena? Naunyn Schmiedebergs Arch. Pharmacol. **353** (1996). https://doi.org/10.1007/BF00169174
53. Steketee, J.D.: Neurotransmitter systems of the medial prefrontal cortex: potential role in sensitization to psychostimulants. Brain Res. Rev. **41** (2003)
54. Vanderschure, L.J.M.J., Kalivas, P.W.: Alterations in dopaminergic and glutamatergic transmission in the induction and expression of behavioral sensitization: A critical review of preclinical studies. Psychopharmacol. (Berl) **151** (2000)
55. Hwan, S.J., Ji, Y.K., Sang, H.K., Lee, M.G.: Role of dopamine receptors on electroencephalographic changes produced by repetitive apomorphine treatments in rats. Korean J. Physiol. Pharmacol. **13** (2009). https://doi.org/10.4196/kjpp.2009.13.3.147

A Study on Methods of Synchronization Between Gestural-Visual and Audiovisual Communication

Richelieu R. A. Costa(✉), Raoni Kulesza, Rafael M. T. Nobrega, Daniel C. França, Tiago M. U. Araújo(✉), Rostand E. O. Costa(✉), and Guido L. S. Filho(✉)

PPGI, UFPB, Scouts Street, João Pessoa, PB 58055-000, Brazil
{richelieu.costa,raoni,rafaeltoscano,tiagomaritan, rostand,guido}@lavid.ufpb.br, daniel.franca@ci.ufpb.br

Abstract. The provision of a window with 3D avatars rendering the machine translation of audiovisual content into sign languages allows for a more significant offer of accessible programming, both on live television and on-demand video services, and reduces the operational costs involved, especially with interpreters. However, additional challenges involve synchronizing the original audio with the equivalent signal presentation, which is slower than human speech. Studies indicate that the presentation duration of a sign in American Sign Language (ASL) is about twice the duration of a monosyllabic word. This work aims to investigate methods and strategies to improve the synchronization of signs with the associated video content. One of the approaches considered is the application of automatic sentence summarization (or compression), which aims to produce a shorter paraphrase for a given input sentence in a simplification process to require fewer signs to convey the original content. Another applicable technique is called elastic adjustment, which consists of modifying the display duration of a media object. Elastic tuning algorithms can adjust the audio speed by up to 10%, up or down, keeping the audio perception quality within acceptable limits. The primary hypothesis is that combining such methods reduces the time required for interpreting content in Brazilian Sign Language, thereby reducing the information desynchronization and prioritizing the most crucial information related to that content. Two experiments with automatic summarization of a news article were conducted. In both, it was possible to display the signing in Brazilian Sign Language using a 3D avatar within the duration of the news segment. The first experiment had a word summarization rate of 31.5%, and the second had a rate of 56.5%.

Keywords: Sign language · Video stream · 3D Avatar Synchronization Strategies · Accessibility

R. Kulesza, R. M. T. Nobrega, D. C. França, T. M. U. Araújo, R. E. O. Costa and G. L. S. Filho—These authors contributed equally to this work.

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, pp. 134–144, 2024.
https://doi.org/10.1007/978-3-031-70439-0_9

1 Introduction

The television (TV) plays an important role in Brazilian society. Besides providing individual and collective entertainment, it also acts as an informative, educational, and social interaction element [1,2]. According to data released by IBGE in 2021 [3], here are 2.3 million people with some degree of hearing impairment in Brazil. The World Health Organization states that 1.5 billion people have some degree of hearing impairment (Deafness) worldwide, and by 2050, it is expected that 1 in every 4 people will have a hearing impairment. [4]. For these individuals, it is necessary that information be available through sign languages[1], which are their natural means of communication.

On the other hand, support for sign languages is rarely explored in video streaming technologies. In TV, for example, support for sign languages is generally limited to a window with a sign language interpreter, presented alongside the original video of the program, recorded live. This solution incurs high operational costs for content generation and production (cameras, studio, staff, etc.), requires full-time human interpreters if sign language is inserted into generated content, thereby restricting its use to a small portion of the programming. In video-on-demand streaming platforms, there is no exploration of features that allow for the inclusion of gestural-visual modality languages. These difficulties result in a significant barrier to communication with others, accessing information, acquiring knowledge, consuming entertainment, among other things.

In 2012, an experiment was conducted with Deaf users to evaluate their level of comprehension of videos with subtitles and with translation into Brazilian Sign Language (LIBRAS) [5]. The experiment took place with 20 Deaf users at the Integrated Support Center for Persons with Disabilities (Funad) in the state of Paraíba. Participants were selected conveniently from LIBRAS training courses at Funad and randomly divided into two groups of 10 users each. In one stage of the experiment, users were invited to watch audiovisual content with subtitles and audiovisual content with LIBRAS translation. Each content was presented twice, and after each presentation, users were asked to respond to subjective questions related to the content presented. [5] shows that considering all content, the average accuracy rate for users who watched videos with LIBRAS translations using a proposed solution was 79.38% with a standard deviation of 9.34%. In contrast, the average accuracy rate for users who watched videos only with subtitles was 25.63% with a standard deviation of 19.86%. This information demonstrates the impact of a solution on communication with Deaf individuals.

Due to the need for an alternative to mitigate the issues mentioned above, providing an option to use 3D avatars rendering machine translation of audiovisual content into sign languages will allow for a broader offering of accessible programming, reducing operational costs involved, particularly with human interpreters. However, additional challenges are involved in synchronizing the

[1] Sign languages are natural languages of a gestural-visual modality, where gestural refers to the set of manual, bodily, and facial linguistic elements necessary for the articulation and visual-cultural significance of the sign.

content's audio with the presentation of equivalent signs, which is slower than human speech [6]. This work explores strategies and adaptations for displaying the natural gestural-visual modality language (using Brazilian Sign Language) with audiovisual content within the duration of the translated audiovisual content, prioritizing essential parts of the message. The following proposals were considered: 1) Automatic summarizing of the messages to be translated into sign language, extracted from closed captions; 2) Extending the display time of the audiovisual content. Techniques for automatic text summarization and elastic adjustment of media display were studied.

The remainder of the article is organized as follows: Sect. 2 presents the theoretical framework, which discusses the main concepts related to this study, such as LIBRAS, automatic text summarization, elastic adjustment, and the VLibras tool. Section 3 presents the details of the studied techniques. Section 4 presents ongoing work. Finally, Sect. 5 discusses conclusions and proposals for future work.

2 Theoretical Framework

2.1 Sign Language

Sign languages are of great interest to linguists and the Deaf community in general because, although they are conventional communication systems, their physical transmission is very different from the transmission of spoken languages. [7].

Sign language, a complex system of gestural-visual communication, plays a crucial role in facilitating interaction between Deaf individuals and society at large. This form of communication is essential for promoting inclusion and ensuring linguistic accessibility for Deaf people, fully allowing them to participate in social, educational, and professional activities.

Several sign languages exist worldwide, each with its own distinct grammar and linguistic structure. In the Brazilian context, the sign language used is Brazilian Sign Language (Libras) [8].

The representation of sign language in the text is called gloss. Gloss is a form of textual expression that follows the grammatical norms of sign language to make communication more accessible. It aims to facilitate understanding for both fluent users of Libras and interpreters and enable translation and textual communication among sign language speakers [8].

2.2 Automatic Text Summarization

Automatic text Summarization is a process that aims to reduce the size of a text by summarizing important points in the input text. It can be described as the task of creating a grammatical summary with minimal loss of information [9]. In Summarization research, selecting relevant sentences from source texts and compressing them for inclusion in the summary is common.

As shown in [10], most sentence compression methods utilize the reduction approach by removing elements, making the message shorter.

Text simplification is a strategy whose primary objective is to reduce both vocabulary and syntactic complexity, improving the clarity and comprehensibility of the content.

When simplifying a text, the focus lies on restructuring sentences and substituting complex terms with more accessible equivalents without compromising the essence of the conveyed message. This approach makes the text more understandable for a broader audience and facilitates interpretation for those who may face reading difficulties due to linguistic or cognitive limitations [11].

2.3 Elastic Time Adjustment for Display

Elastic time adjustment for display is a technique used to dynamically adapt the duration of media objects [12].

This technique is especially relevant in audiovisual communication environments, where synchronization between speech and subtitles plays a crucial role in content comprehension. Elastic time adjustment for display enables dynamic and personalized adaptation, considering individual viewer characteristics such as reading skills and cognitive processing. This promotes a more inclusive and accessible viewing experience, catering to the needs of different audiences, including those with visual or cognitive impairments.

In hypermedia presentations, one of the main tasks coordinated by the presentation orchestrator is the synchronization between the various component objects, which can be achieved through elastic adjustment of object display time. This technique can be applied in presentation time to prevent any temporal mismatch caused by transmission and execution environments. It's possible to compress and expand the display time of hypermedia with inter-media synchronization and reference clock reconstruction [13].

Elastic time adjustment for display represents a field in constant evolution, with continuous enhancements in algorithms and implementation techniques. The integration of methods based on artificial intelligence and machine learning has allowed for additional refinements in this area, enabling more precise and efficient adaptation of subtitles to different viewing contexts. These advances contribute to the creation of more sophisticated and accessible captioning systems, expanding the reach and impact of audiovisual communication across various sectors of society.

2.4 VLibras

The VLibras Suite is an open-source set of tools that automatically translate digital content from Portuguese into Libras, making information more accessible for Deaf individuals on computers, mobile devices, TVs, web platforms, among others [14] (Fig. 1).

Fig. 1. Brazilian Government Website - VLibras Suite

The project is the result of a partnership between the Ministry of Management and Innovation in Public Services (MGISP), through the Secretariat of Digital Government (SGD), the Ministry of Human Rights and Citizenship (MDHC), through the National Secretariat for the Rights of Persons with Disabilities (SNDPD), and the Federal University of Paraíba (UFPB), through the Laboratory of Digital Video Applications (LAVID) [14].

The main components of VLibras are:

- VLibras-Plugin and VLibras-Widget: web browser extensions that allow selected texts on web pages to be automatically translated into Libras and played back through a 3D avatar.
- Vibras-Mobile: VLibras client application for mobile devices (compatible with Android and iOS systems).
- VLibras-Desktop: Tool used to automatically translate selected texts in programs running on personal computers into Libras.
- VLibras-Video: A portal that allows translation into Libras of audio tracks and associated video captions. (Currently only available for captioned and on-demand videos, as the time to render the video with the translation is longer than the video time.).

Another part of VLibras is the backend services, called VLibras-services, which perform automatic translation for the other components (or tools) and store 3D animations of Libras signs that are used to render accessible content after translation. Currently, the sign dictionary has approximately 21 thousand signs in Libras, one of the largest databases of this type in the world. (vlibras citation) [14].

There is also a collaborative tool called WikiLibras, which allows volunteers to participate in the process of building and expanding the sign dictionary [15].

The VLibras Suite is an initiative of extreme importance to ensure access to communication and information for individuals with hearing disabilities in the Brazilian context. Serving as a vital tool for the digital inclusion of these individuals, the suite provides a range of substantial benefits. Among these benefits are compliance with Brazilian legislation, ensuring usability and accessibility for all individuals, regardless of the tools used or their individual limitations.

Furthermore, the VLibras Suite contributes to increasing the number of Internet users and to the overall improvement of web access quality, providing information, knowledge, and means of communication and expression equitably for all. This initiative not only aims to improve the quality of life, employability, autonomy, and social inclusion of people with disabilities but also has a broad social impact, aiming to definitively address the issue of digital accessibility in the country for this population segment.

3 Detailing the Studied Techniques

This article focuses on the proposal of the research group from the Federal University of Paraíba (UFPB) regarding the accessibility barrier still present in video streams due to a lack of synchronization between audiovisual and gestural-visual content.

3.1 Automatic Text Summarization

The text summarization process consists of extracting the main parts of a text by reducing the number of words used to convey such parts. A sentence compression algorithm in Portuguese was proposed in [10], using various machine learning methods.

Two datasets were used: pairs of sentences with long (original) sentences and their respective shortened versions. The approach involves considering that a sentence sr is a possible reduction of another sentence sl when sr can be produced by deleting zero or more tokens from sl. Thus, it is possible to have some pairs with identical sentences (sr is equal to sl) in datasets. The authors highlight the importance of including these cases to have examples with sentences that should not be compressed (since, in machine learning, negative examples are as important as positive ones).

Sentence compression in Portuguese was treated using deletion as a machine-learning classification task. Given a sentence s with a sequence of tokens (including punctuation marks) t1, t2, ..., tn, the goal is to build a classifier that answers

"yes" or "no" to the following question for each token t ϵ s: "Should we remove this token t from s?". Thus, each token t in a given sentence s in our datasets is considered an instance.

The results shown by [10] indicate that the proposal proved more efficient than existing approaches for the Portuguese language. Due to the unavailability of the proposed method using the repository provided by the authors and considering the advancements in Large Language Model (LLM) architectures, we chose ChatGPT4 for caption summarization, which is commonly available for all audiovisual content, to perform translation into sign language then, generating glosses in Libras. [16] provides an analysis focused on the syntactic evaluation of the summaries generated with LLMs. [17] demonstrates that summarizations using LLMs, despite significant stylistic differences such as the number of paraphrases, were deemed equivalent to human-written summaries.

3.2 Elastic Adjustment

A compression and expansion algorithm for compressed audio streams was proposed by [18]. The proposed technique for compression and expansion of MPEG video streams employs inserting and discarding MPEG frames. The standard used in the works is MPEG-2, described by the ISO/IEC 13818 specifications.

For video stream expansion, the algorithm proposed by [13] replicates frames from the input stream according to the presentation order of the frames obtained through the analysis of each frame's type and the storage of I or P frames that were used as a reference in the encoding of other frames. The algorithm then evaluates the timing at which these frames will be presented so that their replication is inserted in the correct order within the output stream.

In the audio compression proposed by [18], the algorithms can adjust the audio duration by up to 10% while maintaining the perceived audio quality within acceptable limits.

3.3 Combination of Techniques: On-Demand Scenario

Given that sign language is slower than human speech [6], it is possible, in an on-demand scenario, to combine the presented strategies to maintain synchronization between the audiovisual and gestural-visual aspects. In this scenario, live content consumption would not be necessary, allowing for the possibility of extending media playback time. By applying automatic text summarization to the subtitles of audiovisual content to extract information succinctly and generate glosses from it, these glosses can be rendered using the sign synthesis component (player) with VLibras' 3D avatars. This approach can solve the synchronization problem and explore a mechanism that inserts sign language, making the content fully accessible to the Deaf community as long as subtitles are available.

This solution can also be used with a playback speed adjustment for the sign language *player*, speeding up the avatar's gestures within an acceptable limit for Deaf comprehension.

In principle, elastic time adjustment can be used for display only with on-demand content, where it is possible to spend more time emptying the buffer without resulting in content loss. In a live TV scenario, however, it would not be possible to perform elastic adjustment at the reception moment to display the media for a longer time due to the occupation of the video buffer.

4 Works in Progress

There is a proof-of-concept application being developed as part of a development phase for the new standard of the Brazilian Terrestrial Digital TV System (SBTVD) [19] (Figs. 2 and 3).

Fig. 2. Window with LIBRAS interpreter avatar on TV

The work addresses accessibility requirements that are being studied and developed, more specifically, support for closed captions, support for sign language, and support for audio description (Fig. 4).

The proposal involves making the LIBRAS window available using 3D avatars and automatic translation into sign language, aiming to reduce operational costs and enable a greater offering of content with sign language in the programming of TV 3.0 in Brazil. In this solution, the broadcaster will transmit a stream containing a sequence of glosses (textual representation in sign language grammar), which will be converted into Brazilian Sign Language (LIBRAS) on the TV itself or on a second-screen device using a *player* and 3D avatars.

For the evaluation of the techniques proposed in this work, a journalistic video segment lasting 2 min and 16 s was selected for testing purposes. The results will be explained below and can be seen in Table 1.

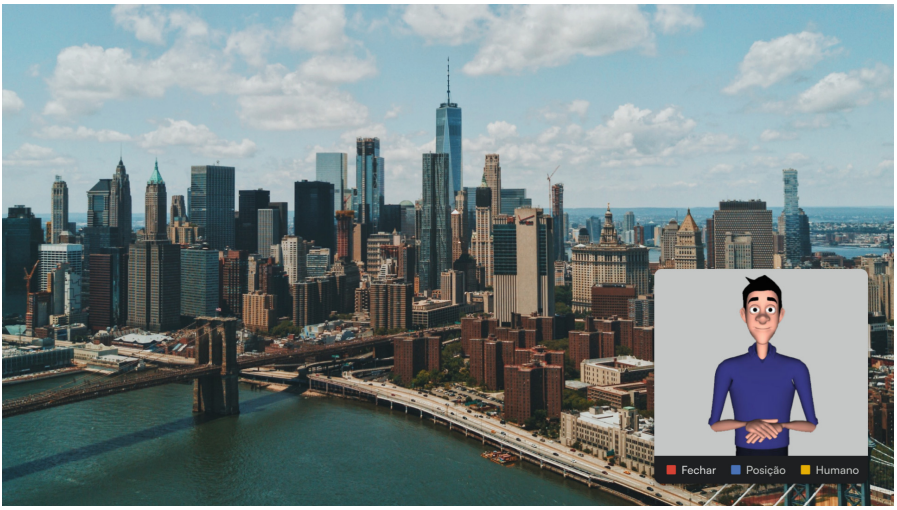


Fig. 3. Possibility of display position adjustment.

Fig. 4. Window with LIBRAS interpreter avatar on a device considered a second screen.

The entire video segment's content was transcribed into text format in captions to generate two summaries of the information, one more concise than the other. Subsequently, the three versions of the text were each translated into glosses to generate a video with a virtual sign language interpreter.

The first text, which is the original, contains 1789 characters with 292 words. From it, a video with a virtual sign language interpreter was generated, with a duration of 2 min and 45 s, 29 s longer than the original video duration. The second text, an automatic summarization, contains 1315 characters, with 200 words. From it, a video with a virtual sign language interpreter was generated, with a duration of 2 min and 26 s, 10 s longer than the original video duration. The third text, an automatic summarization of the first summarization, contains 908 characters, with 127 words, and generated a video with a virtual sign language interpreter lasting 1 min and 9 s, 7 s shorter than the original video duration.

Table 1. Comparison of Caption Summarization

Text	Characters	Words	Duration of Sign Language Video	Word Summarization Rate
Original	1789	292	2 m 45 s	0%
First S.	1315	200	2 m 26 s	31.5068%
Second S.	908	127	2 m 00 s	56.5068%

Thus, considering a 10% adjustment in the duration of the audiovisual content, it is possible to increase the segment's duration by 13.6 s, allowing all the content of the first summarization to be displayed in sign language. The content could be displayed in sign language without altering the video duration when we used the second summarization. Therefore, automatic text summarization of captions is a pathway to ensuring inclusion with accessibility in visual-gestural languages.

5 Conclusions and Proposals for Future Work

The present study analyzed techniques for displaying video content with translation into sign language.

In our preliminary results, it was possible to display a report translated into Brazilian Sign Language within the time frame of the report video. Depending on the compression rate of the automatic summarization of subtitle texts, it is possible to display video content with translation into Brazilian Sign Language without adjusting the subtitled video. A journalistic video segment, lasting 2 min and 16 s, was selected for testing purposes. The preliminary results are at Table 1.

In future work, we aim to recruit Deaf volunteers to evaluate their consumption experience of content. For this stage, we plan to conduct the experiment through a university TV channel, providing accessibility in Brazilian Sign Language and selecting specific test times. Both human interpreters and the 3D avatar from Vlibras will be used for the experiment, with the same glosses extracted from the caption summarization.

References

1. Tavares, T.A., Santos, C.A.S., Assis, T.R., Pinho, C.B.B., Carvalho, G.M., Costa, C.S.: A tv digital interativa como ferramenta de apoio à educação infantil. Revista Brasileira de Informática na Educação **15**(2), 31–44 (2007)
2. Piovesan: Vídeo e tv na educação. Comunicação & Educação **1**, 105–112 (1994)
3. IBGE: Pesquisa Nacional de Saúde 2019: País tem 17,3 milhões de pessoas com algum tipo de deficiência. Instituto Brasileiro de Geografia e Estatística - Rio de Janeiro (2021)
4. OMS: WHO: 1 in 4 people projected to have hearing problems by 2050. Organização Mundial de Saúde (2021)
5. Araújo, T.: Uma solução para geração automática de trilhas em língua brasileira de sinais em conteúdos multimídia. Ph.D. thesis (2012)

6. Bellugi, U., Fischer, S.: A comparison of sign language and spoken language. Cognition **1**(2–3), 173–200 (1972). https://doi.org/10.1016/0010-0277(72)90018-2
7. Sandler, W., Lillo-Martin, D.: Sign Language and Linguistic Universals. Cambridge University Press, Cambridge (2006)
8. Góes, M.C.R.: Linguagem, Surdez e Educação. Autores Associados (2020)
9. Cohn, T., Lapata, M.: Sentence compression beyond word deletion. In: Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008), pp. 137–144 (2008)
10. Nóbrega, F.A.A., Jorge, A.M., Brazdil, P., Pardo, T.A.S.: Sentence compression for Portuguese. In: Quaresma, P., Vieira, R., Aluísio, S., Moniz, H., Batista, F., Gonçalves, T. (eds.) PROPOR 2020. LNCS (LNAI), vol. 12037, pp. 270–280. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-41505-1_26
11. Napoles, C., Van Durme, B., Callison-Burch, C.: Evaluating sentence compression: pitfalls and suggested remedies. In: Proceedings of the Workshop on Monolingual Text-to-Text Generation, pp. 91–97 (2011)
12. Bachelet, B., Mahey, P., Rodrigues, R., Soares, L.F.: Elastic time computation in QoS-driven hypermedia presentations. Multimedia Syst. **12**(6), 461–478 (2007)
13. Alves Cavendish, S.: Algoritmo de Ajuste Elástico Para Vídeo em Fluxos MPEG-2. Ph.D. thesis (2015)
14. Paraíba (UFPB), U.F.: VLibras - Governo Digital (2023). https://vlibras.gov.br/. Accessed 30 Aug 2023
15. Nobre, D.A., Ferreira, M., Araújo, T.M.U.D., Nascimento, I.R., Carvalho, P., Lemos Filho, G.: Wikilibras: collaborative construction of a multimedia dictionary for Brazilian sign language. In: Proceedings of the 17th Brazilian Symposium on Multimedia and the Web on Brazilian Symposium on Multimedia and the Web-Volume 1, pp. 244–251 (2011)
16. Akkasi, A., Fraser, K., Komeili, M.: Reference-free summarization evaluation with large language models. In: Deutsch, D., et al. (eds.) Proceedings of the 4th Workshop on Evaluation and Comparison of NLP Systems, Bali, Indonesia, pp. 193–201. Association for Computational Linguistics (2023). https://doi.org/10.18653/v1/2023.eval4nlp-1.16. https://aclanthology.org/2023.eval4nlp-1.16
17. Zhang, T., Ladhak, F., Durmus, E., Liang, P., McKeown, K., Hashimoto, T.B.: Benchmarking large language models for news summarization. Trans. Assoc. Comput. Linguist. **12**, 39–57 (2024). https://doi.org/10.1162/tacl_a_00632. https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00632/2325685/tacl_a_00632.pdf
18. Maranhao, S.M.D.B.: Ajuste elástico em tempo de exibiÇÃo para fluxos de áudio comprimido. Ph.D. thesis (2006)
19. Costa, R., et al.: Acessibilidade na tv 3.0 brasileira a partir de mídias de legenda, glosa e áudio descrição. In: Anais Estendidos do XXIX Simpósio Brasileiro de Sistemas Multimídia e Web, Porto Alegre, RS, Brasil, pp. 123–129. SBC (2023). https://doi.org/10.5753/webmedia_estendido.2023.236168. https://sol.sbc.org.br/index.php/webmedia_estendido/article/view/25664

Author Index

M. J. Abásolo et al. (Eds.): jAUTI 2023, CCIS 2140, p. 145, 2024.
https://doi.org/10.1007/978-3-031-70439-0