

JAUTI22

XI Jornadas Iberoamericanas sobre Aplicaciones y Usabilidad de la TV Interactiva

NOVIEMBRE

17, 18

2022
Córdoba
Spain

Proceedings
Anales

Editores
María José Abasolo
Carlos de Castro
Gonzalo Olmedo



UNIVERSIDAD
DE
CÓRDOBA

REDAUTI

FREE RED ESPECIAL
FUNDACION

MEDIALABTV
Andalucía

CIT@C
UNIVERSIDAD DE CORDOBA

Springer



diada

JAUTI22. XI Jornadas Iberoamericanas sobre Aplicaciones y Usabilidad de la TV Interactiva.

UCOPress Editorial de la Universidad de Córdoba, España, 2024

Eds. M.J. Abásolo, C. De Castro Lozano, G. Olmedo Cifuentes

ISBN: 978-84-9927-801-8.

1 Preface |

2 Committees |

5 Content Creation and Interaction |

6 Media multitasking as an interactivity factor in the design of post-Covid-19 confinement

Luis Enrique Ibarra

15 Cross-media environment based on iTV to generate personalized Web-based contents as additional information to documentary videos

Alcina Prata, Teresa Chambel, Jorge Ferraz de Abreu

28 A model for digital co-creation and visualizations of collective memories

Ana Velhinho, Pedro Almeida

35 Conformación dinámica de visitas virtuales con vista panorámica

Carlos Alberto Flores Sanjurjo, Joaquín Danilo Pina Amargós, Ariel Alfonso Fernández Santana

50 Audiovisual Consumption |

51 GoWatch: an ITV application to track audiovisual consumption across various streaming platforms

Antoni Oliver, Antoni Bibiloni, Andrea Morey

57 Revisão sistemática da literatura sobre o uso de EEG para mapeamento de emoções durante a fruição audiovisual

Daniel de Queiroz Cavalcanti, Felipe Melo, Thiago Silva, Valdecir Becker

70 Modelos de combinação de serviços Text-To-Speech com locução convencional para a oralização de notícias

Marcelo Alfonso, Pedro Almeida

- 77** **Field trial of an iTV solution to send personalized notifications**
João Encarnação, Ana Velhinho, Simão Bentes, Telmo Silva, Rita Santos
-
- 84** **e-inclusion |**
- 85** **Biblioteca de rutinas de ejercicios con Smart TV para riesgo de caídas en adultos mayores**
Magdalena Rosado, María José Abásolo, Telmo Silva, Stalin Jurado
- 93** **TIC aplicadas a mejorar la enfermedad de Alzheimer de inicio temprano. Análisis sistemático de la literatura**
Ana Camacho, María José Abasolo, Rita Oliveira
- 111** **Técnicas de visión artificial utilizadas en la detección y seguimiento de una pelota en imágenes de un partido de fútbol y su aplicación en un guante háptico para accesibilidad de personas ciegas a la TV**
Diego Villamarín, Lothar Tierra, David Moya, José Manuel Menéndez, Julio Larco
- 129** **TV Notifications: understanding elderly citizens' perceptions of its potential to reduce social isolation**
Juliana Camargo, Telmo Silva, Jorge Ferraz de Abreu
- 137** **Tecnologías exponenciales en el futuro de la salud y la medicina 6P**
Javier Cabo-Salvador, Carlos de Castro, Verónica Cabo, Isabel de Castro, José Aguilar, José Miguel Ramírez, Joaquín Aguilar, Jon Arambarri, Beatriz Sainz, Isabel de la Torre, Johanna Caez, Enrique García
-
- 154** **Digital Infrastructure |**
- 155** **Arquitectura de software para la televisión interactiva adaptada a las condiciones de Cuba**
Ariel Alfonso Fernández Santana, Joaquín Danilo Pina Amargós, Raisa Socorro Llanes
- 168** **Analysis of the value and technological acceptance of OTT**

platforms in Quito, Ecuador

Carina Haro Granizo, Gonzalo Olmedo

**177 Towards on real time early warning system for
microearthquakes at Cotopaxi Volcano**

Román Lara, Santiago Altamirano, Julio Larco, Diego Benítez,
Noel Pérez

**186 Encryption of messages on the transport stream of Digital
Terrestrial Television**

Evelina Silva, Nelson Benavides, Gonzalo Olmedo

**195 Incorporando eye tracking no sistema de visualização de
ondas neurais baseado em EEG**

Matheus Cavalcanti, Felipe Melo, Thiago Silva, Matheus Falcão,
Valdecir Becker

**203 Improvement of network protocol and analysis of security
using aspect of cryptography**

Nisarg Patel, Viral Parekh, Kaushal Jani

Preface

The XI Ibero-American Conference on Applications and Usability of Interactive Digital Television (jAUTI 2022) was jointly organized by CITEC of the University of Córdoba (Spain) and RedAUTI (Thematic Network on Applications and Usability of Interactive Digital Television). This conference took place from November 17 to 18, 2022, in the city of Córdoba (Spain), conducted in an online format. This publication compiles 19 research papers that were presented during the conference, covering various topics such as the design, development, and experiences of applications for interactive digital television and related technologies.

Carlos de Castro

María José Abásolo

Gonzalo Olmedo

Eds.

Committees |

General Chairs

Carlos de Castro Lozano, University of Córdoba, Spain
María José Abásolo Guerrero, III-LIDI –National University of La Plata, Argentina
Gonzalo Olmedo Cifuentes, Universidad de las Fuerzas Armadas ESPE, Ecuador

Scientific Committee

Alcina Prata, Polytechnic Institute of Setúbal, Portugal
Alfonso Infante Moro, University of Huelva, Spain
Ana Martins, Digimedia –University of Aveiro, Portugal
Ana Pisco, Digimedia –University of Aveiro, Portugal
Ana Velhinho, Digimedia –University of Aveiro, Portugal
Anelise Jantsch, Federal University of Rio Grande do Sul, Brazil
Angel Fidalgo Blanco, University of Madrid, Spain
Angel Garcia Crespo, Carlos III University, Spain.
Antoni Bibiloni, LTIM University of Balearic Islands, Spain
Antoni Oliver, LTIM University of Balearic Islands, Spain
Antonio Ocon Carreras, University of Las Palmas, Spain
Armando De Giusti, III-LIDI National University of La Plata, Argentina
Antonio Sagues, University of Córdoba, Spain
Beatriz Sainz de Abajo University of Valladolid, Spain
Bernardo Cardoso, Digimedia –University of Aveiro, Portugal
Carina González González, University of La Laguna, Spain
Carlos de Castro Lozano, University of Córdoba, Spain
Carlos de Castro Burón, University of Sevilla, Spain
Carlos Lazcano, Universidad de la Habana, Cuba
Carlos Matheus, University Carlos III, Spain
Cecilia Sanz, III-LIDI National University of La Plata, Argentina
César Collazos, University of Cauca, Colombia
Cristina Manresa Yee, UGIV-IA University of Balearic Islands, Spain
Cosette Castro, Catholic University of Brasilia, Brazil
David González Avendaño, Smart CiTV, Fundacion Red Especial, Chile
Daniel Gambaro, University of São Paulo, Brazil
Douglas Paredes Marquina, University of Los Andes, Venezuela
Emili Prado, Autonomous University of Barcelona, Spain
Enrique Garcia Salcines, University of Córdoba, Spain
Enrique Rubio Royo, University of Las Palmas, Spain
Esperanza Navarro Pardo, University of Valencia, Spain
Fernanda Chocron, Federal University of Rio Grande do Sul, Brazil
Fernando Boronat, Polytechnic University of Valencia, Spain
Fernando Fuente Alba Cariola, Catholic University of the Holy Conception, Chile
Francisco Alcantud Marin, University of Valencia, Spain
Francisco Montero Simarro, University of Castilla-La Mancha, Spain
Francisco Perales López, UGIV-IA University of Balearic Islands, Spain
Francisco Ruiz Muñoz, University of Huelva, Spain
Gabriel Fernandez, Ramon Llull University, Spain

Gerardo Borroto, University of Córdoba, Spain
Guido Lemos, LAVID – Federal University of Paraíba, Brazil
Gustavo Rossi, National University of La Plata, Argentina
Isabel de Castro Burón, Corporate Communication at IMIBIC, Spain
Israel González Carrasco, University Carlos III, Spain
Ivan Bernal, National Polytechnic School, Ecuador
Joaquín Aguilar, University of Córdoba, Spain
Joaquín Danilo Pina Amargós, CUJAE, Cuba
Johanna Caez, University of Córdoba, Spain
Jordi Belda, Polytechnic University of Valencia, Spain
Jorge Abreu, Digimedia –University of Aveiro, Portugal
Jorge Eduardo Guaman Jaramillo, Technical University of Loja UTPL, Ecuador
José Emilio Guerrero, University of Córdoba, Spain
José Luis Arciniegas Herrera, University of Cauca, Colombia
Josemar Rodrigues de Souza, University of Bahia State, Brazil
José Maria Buades Rubio, UGIV-IA University of Balearic Islands, Spain
Juan Carlos Torres, Technical University of Loja UTPL, Ecuador
Luis Enrique Martínez Martínez, University of Alicante, Spain
Manuel Gonzalez Hidalgo, UGIV-IA University of Balearic Islands, Spain
Marcelo Fernandes de Sousa - Higher Education Institute of Paraíba, Brazil
Mario Montagud-Climent, Centrum Wiskunde & Informatica CWI, Netherlands
Miguel Angel Rodrigo Alonso, University of Córdoba, Spain
Miguel Angel Valero, Polytechnic University of Madrid, Spain
Néstor Daniel González, National University of Quilmes, Argentina
Oscar Mealha, Digimedia –University of Aveiro, Portugal
Patrícia Oliveira, Digimedia – University of Aveiro, Portugal
Pedro Almeida, Digimedia – University of Aveiro, Portugal
Pedro Beça, Digimedia – University of Aveiro
Raisa Socorro Llanes, CUJAE, Cuba
Raoni Kulesza, Federal University of Paraíba, Brazil
Raphael Irerê , Catholic University of Brasilia, Brazil
Rita Oliveira, Digimedia – University of Aveiro, Portugal
Rita Santos, Digimedia – University of Aveiro, Portugal
Roberto Guerrero, National University of San Luis, Argentina
Rostand Costa, LAVID – Federal University of Paraíba, Brazil
Sandra Baldassarri, University of Zaragoza, Spain
Sandra Casas, National University of Southern Patagonia, Argentina
Tania Ribeiro, Digimedia – University of Aveiro, Portugal
Tatiana Tavares, Federal University of Pelotas, Brazil
Telmo Silva, Digimedia – University of Aveiro, Portugal
Teresa Chambel, University of Lisbon, PT
Tiago Maritan, Federal University of Paraíba, Brazil
Vagner Beserra, University of Tarapacá, Chile
Valdecir Becker, Federal University of Paraíba, Brazil

Honor Committee

Manuel Torralbo, Chancellor of the University of Córdoba, Spain

Mohamed Abdelkader

Abdul Basit Syed, Director of the 'WORLD HUMANITARIAN DRIVE', United Kingdom

Ailyn Febles, President of the Union of Informatics of Cuba

Oscar Rueda, Director of the Executive Presidency of the Development Bank of Latin America- CAF. Colombia

Gustavo Santos, Director Regional Department of the Americas UNWTO United Nations Agency, Argentina.

Antonio Santos, President of Tourism and Society Think Tank, Spain

Alvaro Garnica, CEO of Wanderlust Global Life, Spain

Manuel Martín, President of COREMSA Foundation, Spain

Joaquín Ramírez, Secretary General of the Audiovisual Council of Radio Television of Andalucía (RTVA), Spain

Lourdes Arce, Vice Chancellor of Innovation and Transfer of the University of Córdoba, Spain

Lourdes Morales, Deputy Mayor for Digital Transformation of Córdoba City Council, Spain

José Miguel Ramírez, Founding Partner of InteracTVty, Spain

Adrián Fernández, President of Lyons Club Spain

Sebastián Ventura, Vice Chancellor of Digital Transformation and Data Management of the University of Córdoba, Spain

Anabel Carrillo, Director of the QSDglobal Foundation

Paco Lobatón, Vice President of the QSDglobal Foundation

Marisa Vadillo, Director for the DIADA project of the Andalusian Regional Government of the University of Seville, Spain

Francisco Bellido, Head of the COREMSA Chair at the University of Córdoba, Spain

Teresa Muela, General Secretary of the Andalusian Federation of Municipalities and Provinces (FAMP), Spain

Víctor Osuna, Deputy Technical Director of the Digital Agency of Andalucía, Spain

Javier Burón, CEO of Audiense

Pedro Moreno, CEO of MOVATEC

Juan José Rider, Founding Partner of Wul4 and Paythunder

Francisco González, CEO of TRANSEOP

Content Creation and Interaction |

Media multitasking as an interactivity factor in the design of TV content post Covid-19 confinement

Luis Enrique Ibarra¹

¹ Universidad de La Serena, La Serena
libarra@userena.cl

Abstract. *Interactivity*, understood as the potentiality and capacity to offer complementary services in the area of television and as an exchange of information that determines an internal modification of the agents, is the subject for increasing research on specific applications, services and contents connected to society. Since television studies, various perspectives have advanced in explaining both intersubjective and subject-object interaction. The concept of media multitasking can be approached from studies on media effects and from research on mental resources. In this study we have sought an integrated perspective based on a deductive analysis of the motivational desires underlying the choice of modes and contents of participation in relation to core activities such as studying or watching television. We have been interested in think over on the consequences of post-confinement in the acts of choice and use of a cognitive nature that determine simultaneous tasks that require a division of attention resources for interaction with both social and technological systems as antecedents for content design. To do this, we have proposed a statistical study to measure the frequency by which people multitask since the end of confinement and if there is awareness of a choice to carry out another task while rewarding or duty activities at the same time dividing attention. We have arrived at some certainties based on empirical evidence to define design requirements.¹

Keywords: Interactivity, Media Multitasking, Television.

1 Introduction

As a legacy of the 20th century, the first twenty years of the millennium show that interactivity as a potentiality and capacity to offer complementary services in the area of television at a global level has a triple source of technological development and innovation. First, the industry of generic devices and services, economically oriented with a closed structure, dependent on global technological poles; second, the local media and content creation industries with a hybrid economic, social, private or public orientation; and third, research spaces in specific applications, services and content intended

¹ I want to thank PHD Math. Abelardo Araya for his contribution to clarifying and reviewing the statistical procedure applied to data analysis.

2

to cover deep social and cultural problems of democracies, with a high degree of connection with local culture, with an open and collaborative structure. The three sources seem to observe society in its relationship with the media by applying two different approaches to acts of use: on the one hand, physical and behavioral acts, and on the other, the motivating desires underlying the actions and choices about ways, figuration and participation content in relation to the media.

Since television studies, various perspectives from media convergence to multiple screens have advanced in explaining both intersubjective and subject-object interaction. The point of view of our interest is found in social acts of an individual cognitive nature that allow the materialization of simultaneous tasks that require a division of attention resources for interaction with both social and technological systems.

2 Review of interactivity factors

2.1 Operational definition

Interactivity in a broad sense is defined as a quality or condition of *interactive* which in turn is defined as "Action that is exerted reciprocally between two or more objects, people, agents, forces, functions" (RAE, 2022). References to the concept can be found in philosophical and sociological studies. Alexander Thomas Ormond's (1900) concept of interaction can be applied to the behavior of objects and people and involves the internal modifiability of colliding agents. As a consequence of the above, the current of symbolic interactionism accounts for a habitual exchange given by language that determines the construction of meanings of *objects*. Interactivity is understood from the professional area of television as "the ability to offer additional content to the (TV) programs" (Gobierno de España, 2022) and therefore the *potentiality to provide answers*. From the studies and the development of applications for TV, the sense has been expressed through the need to provide on open television, spaces for information, communication, encouragement to citizenship and digital literacy (Castro, 2017), implying an extension of access to information for educational purposes. From this perspective, solutions that incorporate social interaction functionalities can also be considered, for example in the domain of Social TV. (Ferraz de Abreu, Oliveira, & Oliveira, 2020) oriented to the mediated transaction of information between human beings. In relation to the advent of increasingly active audiences, prosumers in an environment *Trans-media Open Online Content* (TOOCs) in playful environments not linked to the game (de Castro Lozano, 2020) they provide us with a type of information exchange leading to new exchanges, facilitated by interconnected systems, active subjects and ideas with educational meaning. From the second screen phenomenon, interactivity can be defined as the use of a device to issue simultaneous feedback to a responsive audience about content received on another device. This leads us to the condition or possibility of *reciprocal action, mediated through devices*, regarding *ideas*, between entities that determines their *internal modifiability*, which implies continuity and simultaneity for learning purposes, therefore, the existence of knowledge threads. A task that emerges from

the above definitions is the enumeration of interactivity variables, such as: subject-object actions, subject-subject actions, object-subject actions, subject choices, perceptual modalities, media modes, devices, media, start-end state, and the ability to multi-task or concurrently process.

2.2 Media multitasking meaning

The term multitask is described from the danger posed by using a mobile phone while driving due to the gradual loss of the ability to maintain simultaneous attention. (Wickens, June 2008). As a derivation, the *Media Multitasking* concept is defined as simultaneously dealing with two or more tasks, one of which corresponds to the *consumption of a medium* (Yoon, Duff, & Bunker, 2021, pág. 1) or also the use of *multiple media* such as *multiscreening* (Segijn, Voorveld, Vandeberg, & Smit, 2017). This has led us to the base question on what foundations allow us to understand the choice of media design, content and connection modes.

3 Operationalizing the concept of media multitasking

3.1 Theoretical background regarding media multitasking

Wang y Tchernev (2012) they guide us on the preponderance of psychological research in mental resources. They point out that *multitasking* behavior in general has been conceptualized under the cognitive framework from the concept of *cognitive demand*—which we can understand as the mental effort required to perform an activity—. Since its origins, multitasking studies indicate that the popularity of media multitasking behavior coincides with the growing evidence of its adverse impact on task performance. (Wang, Irwin, Cooper, & Srivastava, 2015, pág. 102). Various studies have verified the existence of a deficient multitasking capacity of the human being, which has found an explanation from the *Multiple Resources Theory*, inspired according to its author by Kahneman's theory of attention (1973) and in experimental studies of multitasking that have contributed to studies of *divided attention* and the *demand for resources* that are collected indifferently in reserves or pools (Wickens, June 2008). Both limited capacity and multiple resources theory describe the cognitive system as multiple reservoirs of mental resources or energy, each associated with specific modalities of information. (Wang, Irwin, Cooper, & Srivastava, 2015) in this way, visual information can work simultaneously with auditory information. Cognitive resources are used to process information and when they are run up, the probability of learning decreases (Christensen, Bickham, Ross, & Rich, 2015) because the cognitive load increases when performing simultaneous tasks such as watching television and working on a computer, thus probably decreasing the mental processing of TV (Lang, 2000; Christensen et al., 2015)

Authors have wondered what motivates the demand for resources assigned over a period of time to the use of media in conjunction with other activities; what favors persistence and what determines the gradual loss of attention. One answer lies in the

dynamically motivated choice perspective, which posits that "positive feedback originating from underlying motivations reinforces the multitasking behavior system"(Wang & Tchernev, 2012, pág. 494). The tradition of media studies has a good base in the Uses and Gratifications theory, which highlights the pursuit of the achievement of objectives by an active audience, which relates gratifications to media that compete for their attention (Katz, Blumler, & Gurevitch, 1973), also competing with other sources of satisfaction (Otero Bello, 1994). The *offer* of each medium can be defined in relation to its contents, attributes and exposure situations. It is probably not that a medium is permanently eligible, but that its circumstantial contents may be more important when it comes to choosing, maintaining attention and viewing behavior in an interleaved strategy (Segijn, Voorveld, Vandeberg, & Smit, 2017) in the use of resources, as long as the task is not changed for due to tiredness or *boredom* (Yoon, Duff, & Bunker, 2021), the latter that we can operationalize in a continuum up to *entertainment*, a component of an underlying *search for new experiences*, gratification associated with a content that ultimately determines the choice and persistence of a medium and multitasking combination. Interactive services entertainment *offer* may include e-commerce, t-commerce and interactive video (Shin, 2009). In an entertainment-oriented scenario, a transmedia interactive content strategy combines complementary mental resources under serial narratives.

3.2 Post-confinement cognitive acts of choice

The Covid-19 pandemic has configured climate-dependent scenarios that have determined a first confinement in each hemisphere in the coldest months that resulted in a *mediatization* of work and student activities. In April 2020, the number of students forced to stay at home reaches 1.598 billion in 94 countries (Aristovnik, 2020). The confinement seems to have accelerated the virtualization of tasks, it probably meant an increase in autonomy and a consequent decrease in external control, implying an adaptation requirement for many people who evaluated stress as a threat to their pools of resources. (Toh, Ng, & Phoon, 2023). Probably one of the consequences of absence of displacement was having time for highly motivated tasks. Various communication technologies have expanded the possibilities of experiencing immersion by facilitating acts and choices of simultaneous content and media.

Faced with the situation of the probable decrease in infections, we have proposed to understand the *cognitive acts* of choice and use of media tasks by people with the aim of improving the production of relevant content in social coordination and learning that the media can contribute. The strategic design of content and interactive programming from an analysis and disruptive combinatorics for improvement requires creative reasoning relevant to the social context (Frascara, 2022) like being sure if people make strategic choices before planning the use of interactivity media. Design decisions can be defined by applying the *Analytical Hierarchical Process* (AHP) as a procedure to structure hierarchies (Saaty, 1987, pág. 162) to solve the determination of the connection mode decision attributes and the determination of the media (being able to scale to other variables of interest).

4 Empirical study

It has been defined to study the representations of a group of student volunteers, consulting the frequency with which they perform multitasking since the end of the confinement and the return to face-to-face activities, particularly during acts of different motivation: studying and watching TV. We have been interested in measuring the memory fixation of a general idea of the act of multitasking, understanding that the ability to move characteristic of post-confinement can be considered a new, more attractive experience than media. It is thought that students in their decisions differentiate whether or not to multitask when carrying out activities of two types: gratifies or duties.

4.1 Research hypothesis

It is intended to test a measurement procedure for multitasking at present in relation to two categories of tasks and according to the modes of interaction used.

Hypothesis 1 H_{10} : students are not aware that they are multitasking when they study. The responses on the variable *frequency of multitasking with different modes of connection* (x) during the *act of study* (y) do not differ and therefore it is not possible to establish evidence that a pattern of reflexive responses on the behavior of perform multitasking interacting with different devices. In other words, it doesn't matter if their answer to the question of whether they multitask is never or always. The Alternative hypothesis would indicate that there is awareness of a pattern or tendency of action to choose to perform multitasking and therefore intentionality of a rational choice to divide attention. $H_{10}: R_{xy} = 0$; $H_{1a}: R_{xy} \neq 0$

Hypothesis 2 H_{20} : the responses on the variable frequency of multitasking during the act of watching TV did not differ. The Alternative hypothesis would indicate that there is awareness of a pattern or tendency of action to choose to perform other simultaneous tasks. $H_{20}: R_{yz} = 0$; $H_{21}: R_{yz} \neq 0$

4.2 Results of statistical process

A simple data crossing has been carried out that allows visualizing the distribution of the object in its dimensions as a stage prior to correlating the variables.

Table 1 shows that young people consulted are aware that they multitask during the study through different interactive modes, which determine a greater or lesser demand for cognitive resources acting simultaneously where there are complementary perceptual modalities or at least that do not collide, which is basically the basis of the success of the combination between audio and visual.

Table 1: Cross tabulation relative to *Hypothesis 1*
Representation of Multitasking while *Studying* regarding *Connection Mode*

	Other	Mobile + Streaming	Mobile + Tv Set	Pc + Streaming	Smart TV	Total
Never	0	0	0	0	1	1
Rarely	0	1	3	1	0	5
Sometimes	0	1	5	2	0	8
Generally	1	7	3	2	0	13
Always	1	4	4	5	1	15
Total	2	13	15	10	2	42

Variable *multitasking frequency* during the *study*

Table 2: Cross tabulation relative to *Hypothesis 2*
Representation of Multitasking while *watching TV* regarding *Connection Mode*

	Other	Mobile + Streaming	Mobile + Tv Set	Pc + Streaming	Smart TV	Total
Never	0	1	0	1	0	2
Rarely	0	2	0	0	0	2
Sometimes	0	3	2	2	0	7
Generally	0	1	4	3	0	8
Always	2	4	9	4	2	21
Total	2	11	15	10	2	42

Variable *multitasking frequency* when *watching TV*

From the comparison between Table 1 and Table 2, it can be inferred that the use of mobile phones to watch TV content is reduced, probably because the immersion may have been forced by confinement, and released from the obligation with the return to face-to-face, the non-medial environment could be more rewarding. In addition, from the comparison between the 15 responses Always (I multitask when I study) and the 21 responses Always (I multitask when I watch TV) it can be argued that students express a differential between the duty to study and the gratifying activity of watching TV, which implies that the interactive activity would be moderated by the sense of duty that postponing immediate gratification represents for the achievement of productive activity with fruits in the future, linked to the *desire for recognition*.

Table 3: Cross tabulation relative to *Hypothesis 1 Directional measure*
Representation of *TV multitasking* regarding *Connection Mode*

		Value	Asymp. Standard Error	T aproximate	Approximate significance
Lambda	symmetric	.185	.141	1.234	.217
	Connection mode Dependent	.222	.146	1.371	.170
	Multitask frequency while Study Dependent	.148	.167	.823	.410

Variable *frequency of multitasking while studying*

Table 3 indicates that the Multitasking Frequency variable registers a directly proportional correlation with the Connection Mode variable. This finding has allowed us to have a predictive model in the range between .222 and .148

Table 4: Cross tabulation relative to *Hypothesis 1 Symmetric measure*
Representation of *TV multitasking* regarding *Connection Mode*

		Value	Approximate significance
Lambda	Phi	.846	.18
	Cramer's V	.423	.18
		42	

Variable *frequency of multitasking while studying*

Table 4 confirms that the *Multitasking Frequency* variable is correlated with the *Connection Mode* variable.

5 Discussion

From the theoretical perspectives in which this framework has been delimited, it has been determined to establish connections between the traditions of the theory of multiple resources and the theory of uses and gratifications. We are seriously considering *reward identification* as a source of understanding choice and persistence in multitasking.

From the theory of limited capacities, the selection of rationally dosed and combined mental resources during tasks equivalent to duties together with the selection of complementary modes, in the sense that they do not present collision problems that make it difficult to maintain attention, seems to be a response to the question about what foundations allow us to understand the choice of media, content and connection modes.

The collision or complement relationships between different media possibly depend on the one hand on the perceptual modalities connected with modes of connection, and

on the other, on the underlying motivations, which, although dynamic, can be encompassed in a representation or structured idea, which is recovered as evaluative balance of the total experience. These connections can be considered as criteria to define the strategic decisions of media, content and its design. To operationalize the above, various tests have been carried out, arriving at a non-parametric Chi-square test for nominal variables to the variables under study, specifically applying Cramer's V, and a significant correlation of .423 has been detected, which was complemented with a study of Lambda directional measures that has defined a value of .222. This has allowed us to have a predictive model in the range between .222 and .148.

From the use or *utilitarian* perspective, the probable planned choices of devices and means by which interactivity is concretized could be mediated by the importance of the main task, the level of concentration that it requires, the differentiation of mental resources that it demands, the gratification that is obtained from the use of media.

It has been wanted to define a subject matter or research line from the use of media to the bases of the elections and permanence in the simultaneous and intertwined use of media in conditions of security for people. Although studies from different disciplines have been identified, a search has been carried out limited to a theoretical framework whose limit is found in the consideration of at least one medium. However, it is recorded that there are interesting contributions from physiological measurements of mental resources.

Measuring the interactive activity to support a reflection on the gratifications and desires sought by the audiences, allows us to guide decisions on the cross-modal strategy referred to content programming devices. At the same time, certainties based on empirical evidence can be counted on to define design requirements for the content development, thus closing an R&D process. As future work, it is expected to deepen the statistical analysis, as well as its translation into design requirements that support content development initiatives.

References

1. Gobierno de España. (31 de 10 de 2022). Interactividad. *Ministerio de Asuntos Económicos y Transformación Televisión Digital*. Obtenido de Televisión Digital: <https://televisiondigital.mineco.gob.es/>
2. Ferraz de Abreu, J., Oliveira, R., & Oliveira, N. (2020). SecurHome TV: um ecossistema televisivo para a monitorização da atividade doméstica dos seniores. *9ª Jornadas de Aplicaciones y Usabilidade de la Televisión Digital Interactiva*. In: Ferraz de Abreu, J., Abasolo, M.J., Almeida, P., Silva, T. (eds.) CONFERENCE (págs. 73-78). Aveiro: U. de Aveiro.
3. de Castro Lozano, C. e. (2020). TELEDU. Plataforma t-Learning para todos. *9ª Jornadas de Aplicaciones y Usabilidade de la Televisión Digital Interactiva*. In: Ferraz de Abreu, J., Abasolo, M.J., Almeida, P., Silva, T. (eds.) CONFERENCE (págs. 193-199). Aveiro: U. de Aveiro.
4. Castro, C. (2017). El Rol de la Interactividad Hoy en Brasil. *6ª Jornadas de Aplicaciones y Usabilidade de la Televisión Digital Interactiva 6th Iberoamerican Conference on Applications and Usability for Interactive TV jAuti'17*. In: Ferraz de Abreu, J., Abasolo, M.J., Almeida, P., Silva, T. (eds.) CONFERENCE (págs. 142-152). Aveiro: U. de Aveiro.

5. Yoon, G., Duff, B. R., & Bunker, M. P. (2021). Sensation seeking, media multitasking, and social Facebook use. *Social Behavior and Personality, Volume 49, Issue 1, e8918*, 1-8.
6. Wickens, C. D. (June 2008). Multiple Resources and Mental Workload. *HUMAN FACTORS, Vol. 50, No. 3*, 449–455.
7. WANG, Z., IRWIN, M., COOPER, C., & SRIVASTAVA, J. (2015). Multidimensions of Media Multitasking and Adaptive Media Selection. *Human Communication Research*, 102-127.
8. WANG, Z., & TCHERNEV, J. M. (2012). The “Myth” of Media Multitasking: Reciprocal Dynamics of Media Multitasking, Personal Needs, and Gratifications. *Journal of Communication*, 493-513.
9. Katz, E., Blumler, M., & Gurevitch, M. (1973). Uses and gratifications research. *The Public Opinion Quarterly Vol. 37, No. 4*, 509-523.
10. Shin, D. H. (2009). An empirical investigation of a modified technology acceptance model of IPTV. Una investigación empírica de un modelo de aceptación (recepción) tecnológica modificado de la IPTV. *Behaviour & Information Technology*, 28(4), 361-372.
11. Christensen, C. G., Bickham, D., Ross, C., & Rich, M. (2015). Multitasking With Television Among Adolescents. *Journal of Broadcasting & Electronic Media*, 59(1), 130–148.
12. Segijn, C., Voorveld, H., Vandeberg, L., & Smit, E. (2017). The Battle of the Screens: Unraveling Attention Allocation and Memory Effects When Multiscreening. *Human Communication Research*, 295–314.
13. Frascara, J. (2022). Revisiting “Graphic Design: Fine Art or Social Science?”. *She Ji Journal of Design, Economics, and Innovation*, 270-288.
14. Saaty, R. W. (1987). The Analytic Hierarchy Process - What it is and How it is used. *Mat/d Modelling, Vol. 9, No. 3-5*, pp. 161-176.
15. Otero Bello, E. (1994). Teoría de la comunicación y epistemología. *Revista De Sociología*, (9), 15–23
16. Aristovnik, A. e. (2020). Impacts of the COVID-19 Pandemic on Life of Higher Education Students: A Global Perspective. *Sustainability*.
17. Toh, S.-Y., Ng, S.-A., & Phoon, S.-T. (2023). Accentuating technology acceptance among academicians: A conservation of resource perspective in the Malaysian context. *Education and Information Technologies*, 28:2529–254.

Cross-media Environment based on iTV to Generate Personalized Web-based Contents as Additional Information to Documentary Videos

Alcina Prata¹, Teresa Chambel², Jorge Ferraz de Abreu³

¹Superior School of Technology (ESTS), Polytechnic Institute of Setúbal

{alcina.prata@estsetubal.ips.pt}

²Lasige Faculty of Sciences, University of Lisbon

{mtchambel@ciencias.ulisboa.pt}

³Digimedia, Department of Communication and Arts, University of Aveiro

{jfa@ua.pt}

Abstract. Cross-media environments have been gaining space due to their characteristics and capability to support a diversity of contexts of use through flexible solutions and, due to its characteristics, iTV has become an obligatory part of this cross-media environments. This paper briefly addresses the design challenges that need to be considered in the design of cross-media environments capable of generating personalized web-based contents as additional information to video, from different devices. The system designed to illustrate our research, and which evolved from previous versions, is called eiTV and generates a cross-media personalized web-based content, which provides extra information about users' selected topics of interest while watching a specific video from a documentary genre. The web-based content may be generated, accessed, personalized, shared, enriched and (immediately or later) viewed through iTV, PC and mobile devices depending on the users' needs. This paper main focus is on the 'Create functionality' which includes the interfaces designed to generate the web content, and the 'Chat functionality' designed to support users communication needs while watching videos, and enrich the web content. An evaluation, with the participation of 30 elements, from 16 to 52 years old, was carried out with high fidelity prototypes and the achieved results were very optimistic.

Keywords: Cross-media, Transmedia, Video, Web-content, iTV

1 Introduction

Cross-media and transmedia environments, systems and applications have been gaining space in practically all areas due to their characteristics and capability to support a diversity of context of use through flexible solutions [25], [19], [10]. While cross-media systems refer to those where the same message is distributed through different channels/platforms (the keyword is repetition), on transmedia systems, the message is expanded through different devices/platforms (the keyword is expansion) [19],[25]. The success and adoption

of cross-media and transmedia environments is propelled by many factors as the proliferation of new and appealing devices, technological advances, viewers change in technological interests and habits and the systems characteristics, which the most relevant are, flexibility and mobility, so essential to support today's lifestyle [10]. Our world is increasingly cross-media, and one area that benefited from these systems was learning [2, 3, 4, 26], very important considering that, lifelong learning became the new way of life.

When it comes to the medium used to support learning through cross-media systems, video is one of the richest ones. In what relates to the devices used to access videos (as even the TV consumption) it is spread amongst many different devices [7, 15]. However, the devices used to access the video, TV, PC, and mobile devices, are the privileged ones depending, mostly on the age range. Through structure and interaction, these devices can open the door to flexible environments. However, the design of these cross-media environments/systems faces some challenges that may affect their effective use and need to be addressed [12].

The eiTV system designed and developed to illustrate our research, has been through an evolution process of 4 generations of prototypes, all ranging from low to high fidelity prototypes. The fourth-generation prototypes, briefly presented in this paper, were the richer ones in terms of devices and functionalities involved, which increased to match a more flexible perspective. Running from iTV, PC and mobile devices, it provides users with the possibility to choose, from a video, which topics they would want to know more about and with which level of detail. They may decide when and where they would want to access those extra related contents (which we also refer to as informal learning environments) which are generated from iTV, PC and mobile devices, and are presented through the form of an editable and sharable web-based content. The architecture and the main features available in iTV, PC and mobile contexts were already explored and described in previous publications [1],[2],[3],[4]. After the excellent results achieved with the third-generation prototypes (where the videos were from the well-known series CSI), the fourth generation were conceptualized and developed to provide continuity to the research opportunities identified [4]. From the research opportunities identified, we have considered that, two of them, the use of documentary videos (videos typically more dynamic and that by its nature are typically watched in a lean forward mode) and social communication tools, would have a better chance to succeed if implemented together. In fact, the existence of social communication tools was anchored in the use of documentary videos, resulting in a return of '1+1>2' as we have envisioned and was later confirmed through the evaluation process.

After this introduction, Section 2 includes a review of related work and concepts, Section 3 describes the design challenges of cross-media applications in that context, Section 4 presents some of the most important design decisions, Section 5 describes the evaluation process and, finally, Section 6 presents the conclusions and perspectives for future research and developments.

2 Related Work

This section addresses some of the more relevant related research studies in cross-media environments that include the same or similar devices and/or have informal/formal learning goals and uses integrated social communication tools.

The TAMALLE project [20] developed a 'dual device system' for informal English language learning, based on watching iTV and selecting what to access later on mobile phones. This was an interesting system

capable to accommodate different cognitive modes and different contexts of use, especially, if considering the mobile phone possibilities. Obrist et al. [23] 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 [24] 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 other device. Each device maintains awareness of each other and are 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 contents presented as extra information. Our work is more flexible in these concerns. 2BEON [13], 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 of including a sharing functionality in eITV. Cronkite [18] 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 latter view (and that might be the viewers’ preference). Our application stores the related information for later use, the simultaneous use of iTV and PC is a possibility but not the only option, 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.

3 Cross-Media Design Challenges

This section describes the key aspects, cognitive and affective, that need to be considered to effectively design cross-media environments and interfaces, with a special focus on the design challenges associated with video and different devices.

Media and Cognition: Norman’s view [8] defines two fundamental cognitive modes. The experiential mode allows us to perceive and react to events naturally and without cognition, but require 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.

Cross-media Interaction, Conceptual Model and User Experience: the main challenges of cross-media interaction design described by [17] 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 [9]. 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 [16], 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 [23], 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 Cross-media HCI: In this context, the migration of tasks is supported via cross-media usability and continuity, influencing on how well and smoothly users' skills and experiences are transferred across the different devices [21] 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) [11].

Designing for Different Devices and Contexts of Use: Cross-media design involves designing interfaces for different devices. To understand the devices and have each one doing what it is most suited for, the best approach is usually to study each particular situation, including device characteristics and cognitive and affective aspects associated to its use: why people use them, in which mode, compare them, etc., and the design guidelines for each device [1] followed by an adequate combination.

Supporting Communication tools in Cross-media contexts: The use of communication tools integrated with cross-media 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 Cross-Media Design in eiTV

In brief, this Section presents main functionalities and design options concerning the eiTV Cross-media system, in response to the challenges identified in Section 3. A specific focus was given to the Create and Chat functionalities considering that it refers to what's new on this paper.

4.1 eiTV Architecture

The eiTV system is a portal aggregator of all the functionalities which may be accessed from any of the devices (iTV, PC and mobile phones) thus working as a true ‘ecosystem of devices’. Through the portal we may: generate web contents; see, edit and share web contents, upload files, change profile, etc. In sum, everyone may receive web contents generated by the eiTV, a characteristic that provides **flexibility** to the application.

4.2 Flexible Navigation Model

We opted for a menu style navigation which 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. This 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 **cross-media 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 Eitv Functionalities

a) The **Create functionality** allows users to watch videos and select topics of interest for further information. The selectable topics appear underlined in the subtitles (see Figure 1 a) and for the first time, due to the nature and characteristics of the videos being used (documentary), each line may contain one selectable topic. Let’s imagine the traditional subtitle with 2 lines. The first line selectable topic appears in a bright white color for a few seconds, then it becomes grey (signaling that is no longer selectable, at least at that moment) and the second line topic is the one that becomes bright white. The selected topics change their color to red and a little box with a ✓ inside, appears in front of the line where the selected topic is, in order to create visual feedback on the users selected topics. When the videos being watched don’t have subtitles, only the selectable topic words appear onscreen to let the users know which topics may be selected. As to the information about a topic, three levels were proposed, from less to high informative (as presented in table 1):

Table 1. Levels of information about a topic

Level	Description/Ambit	Interface:
Level 1 (topics)	Only implies the use of the designed solutions in order 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

At any moment, the user can change between levels of information by pressing button 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 the previous prototypes, that they play an important role to accommodate viewers' 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 [6].

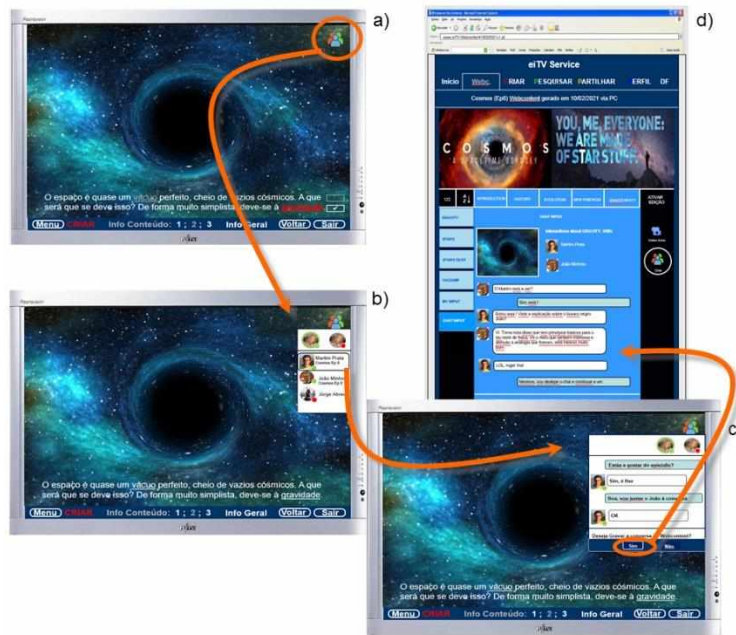


Fig. 1. a) Chat icon on the right corner of the video; b) List of users that use the eiTV system (with the green bullet to show the ones online and the red bullet to show the ones offline); c) Chat messages changed between 2 users and the button to save those messages on the web content visible; d) Chat messages saved on the web content specific ‘CHAT input’ place

As to the personalized generated web-based content, its structure particularities may be seen in table 2.

Table 2. Webcontent structure and characteristics

Name	Description
My input	Each web content is organized as follows. The left side menu contains all the topics selected by the user, presented by the order of selection in the video, to improve contextualization, but the user may choose to see them by alphabetical or logical (content dependent) order (see figure 1 – b and d). This option was designed to take advantage of each device characteristics to provide flexibility.
CHAT input	On the web content left side menu, below the ‘My input’ place there is the ‘CHAT Input’ place (see figure 1 – d) where the chat conversations that took place inside the system are saved and presented (if the user decides to store them). The description of the chat functionality is presented in g). This functionality was made available to provide users with more flexibility and to support their social communication needs without losing the sense of unity (considering that they didn’t need to use external means and devices to communicate, just the ones included

	in the eiTV system in a specific an integrated way). As to the interaction model and interfaces (see Figure 1), they were designed considering the UX and the devices and media characteristics in order have each device contributing with its strongest features.
Editing	Each web content has the possibility to be edited. This edition ranges from uploading textual information (if through the TV set) or textual information and files (if through PC or mobile devices) or GPS coordinates (if through the mobile), to delete, move, edit, import or define privacy status of the web content, a topic of the web content, a category from a specific topic, etc. This provides flexibility, control, consistent interaction and takes advantage of each device characteristics and UX.
Continuity and Contextualization	the WebContent continuity and contextualization, it was supported via the use of some excerpts from the original video, namely, the excerpts that were being watched in the moment of the topic selection.

b) Search functionality allows searching videos based on different criteria. Video criteria: title, actor name, etc; and system criteria: video with or without web content(s) already generated. This provides **flexibility**.

c) The Share functionality is activated only after users accessed the Create or Search functionalities. The share functionality allows sharing the generated web content or retrieved video (with or without web content), with their contacts. On this functionality flexibility and error prevention were improved.

d) The User Profile functionality allows to upload users' personal data from their social network, allows validation, etc. The user profile information is used to personalize the web content, thus improving **flexibility**.

e) The DF functionality was designed to have each device doing what it is most suited for. To achieve this goal, **contexts of use, device characteristics and cognitive and affective aspects** associated to the devices use, were studied. In the case of mobile devices functionalities, the following were made available:

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 devices; contrary to TV and PC, mobile devices may be used everywhere, even when users are standing up, meaning that any extra time may be used.

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);

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

f) The Devices Synchronization functionality - The possibility to synchronize devices was designed and implemented to allow the application to work as a true ecosystem of devices.

g) The Chat functionality – the chat functionality was designed and made available to provide users with more **flexibility** and to support their **social communication needs** (so important on 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) as will be explained next. After logging in, and unless they specifically decide to appear offline, by default, users will appear as online to all the others. To facilitate the learning and use of this functionality, and potentiate the **UX**, references and characteristics common to other similar communication tools were followed. Thus, users may personalize the chat with their pictures (like it happens in WhatsApp, messenger, etc). When online, users have a small green bullet associated with the picture (like on Instagram, etc) and when offline users have a small red bullet associated with the picture (like on TEAMS), Figure 1 - b. There is an icon on the right corner of the video (see Figure 1 – a). When the users choice is to enter in offline mode, this icon will be available onscreen only

for the first 3 minutes of the video, next disappearing until activated again through a specific key on the remote (when watching the video through the TV set), through a simple touch on the video (when watching the video through the smartphone) and through a click/touch on the icon (when watching the video through the PC). At any time, users may see the list of other users that also use the system, choose one or more, and start interaction through chat conversation. The messages appear on the right corner of the screen (Figure 1 – c) with the indentation typically used on other communication tools (the sent messages appear aligned to the right while the received messages appear aligned to the left and with the user picture, similarly to what happens with the messenger, Instagram, WhatsApp, etc). These design options help reduce the learning curve and take the best advantage of the UX. Important to refer that when the user is watching the video through the TV and activates the chat option, the devices synchronization functionality (already described in f)) is automatically activated and synchronizes the TV set with the mobile device. From then, the mobile device is prepared to be used as a 'keyboard' to the messages that will appear on the TV screen. This option improves the system **usability** (considering that writing through a TV remote is a very difficult and time-consuming task), takes the best advantage of each **device characteristics** and, once again, the system works in an **integrated** way as a true **ecosystem of devices** providing users with a sense of unity and continuity. When a user receives messages, a 'pause option' automatically appears onscreen. This option gives users the possibility to pause the video, right away or later, and focus on answering the chat messages without losing the video content and the opportunity to choose topics of interest. When the user is watching the video through the other two devices (mobile or PC) and activates the chat option, the devices synchronization is not activated considering that the devices in use have suitable keyboards to write the messages and don't need extra resources to improve usability. Important to note that the chat option was made available only during the visualization of the video, thus giving support to the users' socialization needs. Its usual to see these socialization needs arise during video watching, not only due to the videos dynamic nature, but because socialization it's a phenomenon pretty much associated to the TV consumption which typically occurs in group. It is normal to share and comment with others what we are watching, mainly if it is news, sports or documentary related. When users ask to exit the chat, by default the system asks if they want to save the conversations. If the answer is yes, the messages are stored in the web content, in the 'CHAT input' area (see Figure 1 – d), thus providing users with more flexibility and personalization possibilities. Important to note that the chat may be used in any circumstance while watching a video. However, to save the chat conversations is only possible after choosing the first topic of interest, which is the action that makes possible later generate a related web content.

Consistency in UX and the perception of the system coherent unity independently of the device being used was also a priority. Despite having considered the mobile device characteristics and contexts of use in the design, towards a more simplified design, we decided to keep a coherent layout in terms of colors, symbols and other graphic elements, as navigational buttons, in order to better contextualize users, give them a sense of unity in their UX and to allow a smooth transition among media and devices. This way, it was possible 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 contents, helping to conceptually understand the system as an 'ecosystem of devices'.

5 Evaluation

The UX evaluation is important in any type of system and context [14]. The UX methods and measures considered relevant for this specific case were: observation, case studies, lab experiments, experience sampling method, questionnaire, interviews and focus groups. In what relates to the design of the new functionalities and interfaces, 3 usability experts were consulted. As to the final evaluation, which occurred from February to November 2021, there were 30 participants, ranging from 16 to 52 years old, which were grouped into 2 evaluation groups: Group 1 (G1) composed of 15 participants, aged between 18 and 52 who already participated on previous evaluations and Group 2 (G2) composed of 15 participants, aged between 16 and 43 who never were 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. Each subgroup was composed of 1 person with less than 25 years old, 2 persons between 25 and 40, and 2 persons with more than 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 evaluators that already participated on previous evaluations was to understand to what extent this system resulted more complex and/or hard to use when compared with the previous versions. The idea of using a group of evaluators that never interacted with the system was to try 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. The evaluation process started with a demonstration of the high-fidelity prototype using the three devices involved (iTV, PC and mobile) and the functionalities being tested. Then, users were asked to perform tasks that allowed us to evaluate:

- a) the specific interfaces created for choosing topics of interest, from the three devices, while watching documentary videos;
- b) the communication tool (chat) functionality interfaces designed for the three devices, and this functionality interest.

The mentioned tasks were performed in three different contextual scenarios using three different devices: at a simulated 'living room', each user used the iTV to visualize the documentary videos, generate a web content and used the chat to communicate with other testers that were online (each user was in the living room alone). Next, users went to the library which, although surrounded by people, is a quiet place and a context that replicates the scenario of a medical waiting room. Here they used the PC to watch new documentary videos, generate a web content and use the chat. Next, they moved to the school bar and repeated the previous tasks but through the mobile device. Finally, users were asked to fill a questionnaire and were interviewed. The questionnaire was based on the USE questionnaire (usefulness, satisfaction, and ease of use) [5]; the NASA TLX questionnaire (cognitive overload) [22]; and usability heuristics.

Independently of the group, medium and high technological literacy categories reacted well to difficulties. However, when considering low technological literacy categories, it was possible to see that, as expected, in the presence of difficulties, the 2 older participants from G2 reacted with higher resistance and discouragement than G1. From Table 3 it is possible to see that, for both groups, the PC is the preferred device to generate the web content (G1 – 40%; G2 – 47%). As to the second preferred device, G1 chose mobile (33%) while G2

chose iTV (33%). These results were somehow unexpected but, from the observation and qualitative evaluation it was possible to perceive that both groups considered the iTV interfaces well designed and intuitive. G2 chose the iTV as their second preferred device to choose topics of interest (33%), instead of the mobile (contrary to what happened with G1). This was more visible in the categories with poor and medium technological literacy. It makes sense considering that even with the need to back forward some videos they feel more comfortable using iTV than mobile devices. Anyway, this is a good indicator considering that G2 was testing the prototype for the first time and with a higher level of interface complexity when compared to previous evaluations. This indicates that the iTV interaction was well designed and that users are becoming more comfortable in what relates to interact with the TV set. However, the major critic was about the difficulty that they felt in choosing topics of interest on time, mainly when from iTV. In fact, now there are 2 topics on each ‘subtitle sequence’ so the time to choose a topic was reduced to half (when compared to previous scenarios when there was only one eligible topic). The problem was not the increased complexity itself, not even the selected topics feedback (which was considered good through all devices), but the necessity of a faster reaction and response. Some users had to pause the video and back forward to be able to select the topic on interest. That solution worked but, with some level of frustration when repeated several times. Thus, something needs to be done to solve this constraint, namely, rethink and improve the interfaces and also consider and explore new ways of input.

Table 3. Preferred Device to Generate the Web Content

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

In terms of the chat functionality, something that was implemented on mid fidelity prototypes for the first time and was never tested before, both groups reacted well as may be seen from table 4. However, from the qualitative evaluation it was possible to understand the weak results obtained in relation to the easiness of use. In fact, what was harder for them was the pitch used in the chat, which was considered very small to their necessities. When this option was designed, we kept the chat in a small part of the top right side of the screen in order to avoid losing the connection with the video being watched. In fact, users could be more interested in continuing the viewing process for a while, or even until the end of the video, before answering the chat. The design of the solution needs to be adapted to improve its usability.

Table 4. Evaluation of specific mobile devices functionalities

Chat functionality:	G1	G2
Useful	15 (100%)	13 (87%)
Easy to use	11 (73%)	9 (60%)
Easy to learn	14 (93%)	13 (87%)
Like to have it	12 (80%)	13 (87%)
Recommend to a friend	13 (87%)	14 (93%)

The preferred device to use the chat was the mobile (somehow expected), when through the TV the use of the mobile as keyboard was considered interesting (but distractor from the videos being watched). As to the intention of transmitting a sense of unity through the chat, it was achieved with success (G1: 100%; G2: 93%).

As a whole, the eITV cross-media system was evaluated and the results are presented on table 5.

Table 5. Overall Evaluation of the Whole eiTV Cross-media

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

As can be seen, the evaluation of groups G1 and G2 are close which was a very good surprise but, as expected, better in G1. As to the two users, from G2, that didn't find the system useful, there was a 16 years boy with high technological literacy that didn't considered the system useful mainly due to the fact that he doesn't appreciate documentaries and, even if so, he considered that the system, in spite not difficult to use for him, would be very hard to use for oldest people like his grandparents. The other user, a 42 years old participant with low technological literacy, he didn't find the system useful because he prefers to search things online through his mobile device when needs extra information. In general, the results were considered very good.

Important to refer that educational tests were not carried out. As mentioned previously, the generated web-contentes aggregate additional information based on the viewers 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

The evaluation results were very encouraging. In many aspects, the designed functionalities (Create and Chat) and the system flexibility were perceived as useful and an added value in the cross-media research area. Some design options allowed to accommodate the changes in users' cognitive mode (e.g., information levels), and the prototype was designed and tested in real scenarios and contexts of use. Considering the design framework followed, the trends in the use of multiple devices, and the results of this and previous studies, we have reasons to believe that our goal for this cross-media context is worth pursuing and that we can achieve quite good results with all the devices in different scenarios. However, this study is limited considering that only the Create and Chat functionalities were implemented and tested. Thus, as future work, we intend to improve the interfaces (mainly the iTV ones), the create functionality, the chat interface model, test de system as a whole unit, and continue exploring the devices technological advances to create new input solutions and functionalities capable to better support users needs and different cognitive modes. A continuous improvement of the interfaces, so they may become easier to learn and adopted by an elderly population, is also a goal.

References

1. Alcina Prata and Teresa Chambel. 2011. Going Beyond iTV: Designing Flexible Video-Based Cross-media Interactive Services as Informal Learning Contexts. In Proceedings of 9th European Conference on Interactive TV and Video: Ubiquitous TV (EuroiTV 2011, in coop with ACM, pp. 65-74, Lisbon, Portugal.
2. Alcina Prata and Teresa Chambel. 2019. 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, 59-71, Rio de Janeiro, Brasil.

3. Alcina Prata and Teresa Chambel. 2020. 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. https://doi.org/10.1007/978-3-030-56574-9_5
4. Alcina Prata and Teresa Chambel. 2021. Personalized Interactive Video-based Crossmedia Informal Learning Environments from iTV, PC and Mobile Devices – the Design Challenges. In: Abásolo M. (eds) Applications and Usability of Interactive TV. jAUTI 2020. Communications in Computer and Information Science. Springer, 89-104.
5. Arnold Lund. 2001. Measuring Usability with the USE Questionnaire. Available at: <https://garyperlman.com/quest/quest.cgi?form=USE>, last accessed 2021/06/16.
6. Bernardo Cardoso and Jorge Abreu. 2021. TV Personalisation: Blending Linear and On-demand Content in the Living Room. International Journal of Entertainment Technology and Management, 1(2), 162–177. <https://doi.org/10.1504/IJENTTM.2021.116769>
7. Bernardo Cardoso. 2022. A Unificação no Consumo de Conteúdos Audiovisuais: contributos para a experiência de utilização e sugestões para operadores. PhD thesis, Communications and Arts Department, Aveiro university, 14 January 2022. Portugal.
8. Donald Norman. 1993. Things that Make us Smart. Addison Wesley Publishing Company.
9. Donald Norman. 2002. The Design of Everyday Things. New York: Basic Books.
10. Henry Jenkins. 2009. Transmedia missionaries: Henry Jenkins. Published in 23 of July 2009. Available at: <http://www.youtube.com/watch?v=bhGBfuyN5gg>, last accessed 2020/11/19.
11. Jakob Nielsen. 1989. Coordinating User Interfaces for Consistency. Neuaufgabe 2002 ed., the Morgan Kaufmann Series in Interactive Technologies, San Francisco, CA, USA.
12. Jonathan Taplin. 2011. 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, pp. 9, Lisbon, Portugal.
13. Jorge Abreu. 2007. Design de Serviços e Interfaces num Contexto de Televisão Interactiva. PhD Thesis, Communications and Arts Department, Aveiro University. Portugal.
14. Jorge Abreu, Pedro Almeida and Telmo Silva. 2016. A UX evaluation approach for second-screen applications. In Communications in Computer and Information Science (Vol. 605, pp. 105–120). https://doi.org/10.1007/978-3-319-38907-3_9
15. Jorge Abreu, Pedro Almeida, Ana Velhinho and Enrickson Varsori. 2019. Returning to the TV Screen. In Managing Screen Time in an Online Society. IGI Global, 146–171. <https://doi.org/10.4018/978-1-5225-8163-5.ch007>
16. Katarina Segerståhl and Harry Oinas-Kukkonen. 2007. Distributed User Experience in Persuasive Technology Environments. in: Y. de Kort et al. (Eds.), Lecture notes in Computer Science 4744, Persuasive 2007, Springer-Verlag.
17. Katarina Segerståhl. 2008. Utilization of Pervasive IT Compromised? Understanding the Adoption and Use of a Cross Media System. In Proceedings of 7TH International Conference on Mobile and Ubiquitous Multimedia (MUM'2008) in cooperation with ACM SIGMOBILE, 168-175, Umea, Sweden.
18. Kevin Livingston, Mark Dredze, Kristian Hammond and Larry Birnbaum. 2003. Beyond Broadcast. In Proceedings of ACM IUI'2003, The Seventh International Conference on Intelligent User Interfaces, (Miami, USA, January 12-15, 2003), 260-262.
19. Kevin Moloney. 2014. Multimedia, Crossmedia, Transmedia... What's in a name? Published on 21 of April 2014. Available at: <https://transmediajournalism.org/2014/04/21/multimedia-crossmedia-transmedia-whats-in-a-name/>, last accessed 2020/11/30
20. Lyn Pemberton and Sanaz Fallahkhair. 2005. 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.
21. Murielle Florins and Jean Vanderdonck. 2004. Graceful Degradation of User Interfaces as a Design Method for Multiplatform Systems. Proc. of the ACM International Conference on Intelligent User Interfaces (IUI'04), 140-147,

Funchal, Madeira.

22. NASA. 2019. NASA TLX Paper and Pencil Version. Available at: <https://humansystems.arc.nasa.gov/groups/tlx/tlxpaperpencil.php>, last accessed 2021/06/16.
23. Obrist, M. and 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.
24. Reed Martin and Henry Holtzman. 2010. Newstream. A Multi-Device, Cross-Medium, and Socially Aware Approach to News Content. Proc. of 8th European Interactive TV Conference (EuroiTV 2010), in coop with ACM, 83-90, Tampere, Finland.
25. Renira Gambarato. 2020. Crossmedia, Multimedia and Transmedia. Published on 20 of October 2020. Available at: <https://www.youtube.com/watch?v=G3wdbajO6js> , last accessed 2020/11/30
26. Stefano Bonometti. 2017. Learning in Cross-Media Environment. International Journal of Web-Based Learning and Teaching Technologies, 12(4), 48-57. October 2017. DOI: 10.4018/IJWLTT.2017100105

A model for digital co-creation and visualization of collective memories

Ana Velhinho^[0000-0001-9978-8317] and Pedro Almeida^[0000-0001-5878-3317]

Digimedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal
{ana.velhinho, almeida}@ua.pt

Abstract. In the domain of cultural events, heritage and memory institutions, the role of crowdsourcing and user-generated content keeps growing impact on the emergence of social archives, the enrichment of digital collections and sharing of collective memories and experiences. In this context, a technology-mediated model through a participatory digital platform is proposed based on a research process comprising case studies, prototype testing with users and exploratory interviews with experts. The model combines online and offline strategies, namely visualization, gamification and locative practices, to potentiate the aggregation, correlation and co-creation of collaborative digital resources. Hence, the model is presented through a set of operative components and principles, to guide the development of prototypes for specific contexts and territories. The participatory nature of this approach intends to empower individuals, municipalities and cultural entities like museums, libraries, and local organizations to create and expand their community archives around past and current events, by presenting them in a multidimensional way. The receptivity and recommendations from potential users and experts – regarding participation, mediation, and visualization aspects – allowed us to improve the MNEMOSCOPE model and understand how to adapt the participatory platform to specific needs. Hence, the development and testing of a functional platform will be held within an R&D project, focused on cultural events and experiences in Portugal.

Keywords: Collective memory; Participatory platform; Digital archives; Co-creation.

1 Introduction and methodological approach

The current investigation is framed on Participatory Culture [1] in which digital storytelling [2,3] and interactive visualizations [4,5,6], driven by social computing and by participatory and co-creation methodologies [7], support not only the entertainment and creativity domains but also cultural and knowledge production [8].

Lev Manovich [5] described the impact of new media in the 21st century as a computerization of culture, which redefined the mediascape and introduced new cultural models and forms such as the database. This paradigm shift influenced the whole media cycle, further propelled by Web 2.0, facilitating content flow and production, which are the core of Convergence Culture [1]. Consequently, users acquire an enhanced role as *prosumers*, as they are not exclusively consumers but also

contribute with their content – known as User Generated Content (UGC). In this sense, convergence is also profoundly linked to participation and building a sense of community, as people actively create and share content around a common interest, relying on a group effort.

Currently, social dynamics among online communities are intrinsic to participatory societies, increasing possibilities for inclusion through UGC and life testimonials, which are gradually becoming relevant contributions. Given the acceleration of the ongoing process of culture digitization, the construction of collective memory is significantly affected by the forms and technologies of capturing and accessing records of those memories [9,10], which also impacts the internal workflow of memory institutions (archives, libraries, and museums).

On the one hand, this shift also influences the research field giving rise to emerging knowledge areas supported by computing and digital data processing (including big data) to study cultural processes and artefacts, namely the Digital Humanities [11] and the Cultural Analytics [12]. On the other hand, the popularization of interactive visualizations opens new possibilities for visual presentation, since the same data can be tailored to tell different stories [3,6], providing users more control to explore, co-create, and convey group storytelling [2]. In this way, digital platforms offering exploratory interfaces [4,13,14] propelled by the self-organized crowd [15,16] may motivate the sharing of multiple points of view for the digital enhancement of experiences as collective constructs.

In this context, through the proposal of a model mediated by a participatory digital platform, the present investigation aimed to encourage networked co-creation through the combination of digital resources to contribute to enhancing shared experiences within the domains of cultural events and intangible cultural heritage, partially dispersed or with limited access. Also, for visualizing real-time experiences in a multidimensional way and preserving them for future memory.

The concept of the participatory platform supported by the model is giving users the autonomy to capture and share records of collective experiences to promote multiple narratives from correlated views (e.g., based on time, geolocation, semantic correlation, networks and links between users and content, etc.). This platform has the following objectives: to strengthen the participatory culture by giving people a voice through UGC; bringing cultural institutions and local associations closer to their audiences; enhance the way we interact with places and in/with communities; aggregate and enrich multimedia records of locative experiences; explore the diversity and creativity of multiple perspectives of collective memories; and co-create a living archive based on paradigms of our time.

Regarding the methodology, the investigation [17] developed an empirical study with an exploratory and qualitative approach centred on the Grounded Theory [18], which feeds the subsequent stages with successive data collection and systematization. The empirical study [19] adopted a combination of methods and unfolded in three stages:

- 1) *Study of multiple cases of participatory projects and apps with a focus on visualization*: which allowed the identification of relevant interface and visualization models for the platform, as well as their systematization through the

creation of a visualization taxonomy and a participation taxonomy for the analysis and design of participatory projects.

- 2) *Semi-functional prototyping of a mobile application*: which allowed understanding the dynamics of a digital community around visual records of musical events, to obtain opinions on motivations and preferences of a young target audience regarding locative content sharing.
- 3) *Interviews with experts*: This allowed assessing the receptivity to the platform's concept by experts and cultural institutions that promote events, leading to the identification of relevant contexts of application, as well as strategies, opportunities, and challenges that were paramount for the design of the model.

The modelling process goes through a theoretical operationalization that guides prototyping and knowledge transfer, with the flexibility to allow adaptation to different contexts [20]. This investigation, aimed to identify and streamline participation mechanisms to clarify the relevance of the participants' contributions to a collective interest, through visualizations within an iterative dynamic of knowledge discovery and creation [21], carried out together with people through digital and face-to-face mediation dynamics [7,22]. Thus, the physical experience of events reflects on the remembrance mechanism of the digital platform, which also intends to enhance the face-to-face experiences as an expanded hybrid reality based on a permanent process of content activation using mediation strategies.

In terms of structure, the document comprises three sections: the first frames the context, objectives, and methodological approach of the investigation; the second presents the model and its components and guiding principles; and the third systematizes the research contributions and its continuity through future work.

2 The MNEMOSCOPE model

The technology-mediated model MNEMOSCOPE aims to represent multiple points of view on collective memories of experiences in their observable and sharable dimensions using a participatory platform. Hence, it works, not as a search engine or repository based on taxonomic inventory but, as a storytelling system and a trigger of remembrance. Based on a cyclical and non-sequential collaborative process (Figure 1), the digital platform underlying the model aims to explore and discover content through association and derivation. Additionally, stimulated by correlated visualizations of shared records, the model combines a cognitive dimension of content enrichment with an emotional dimension based on the community feeling of belonging to encourage participation.

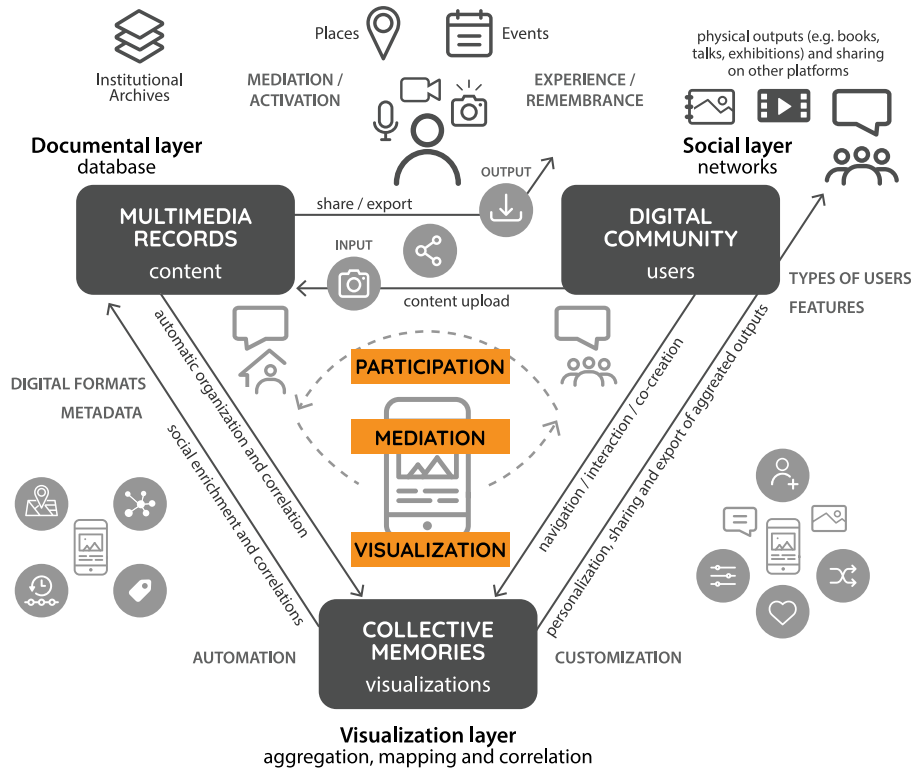


Fig. 1. MNEMOSCOPE model for digital co-creation and visualization of collective memories based on a participatory platform.

Figure 1 systematizes the model elements, which materialize a hybrid experience, resulting from the synergy between the physical instance (the places where events and experiences occur) and the digital instance (the technological mediation of the platform), taking advantage of georeferencing and locative dynamics during face-to-face participation. Thus, the model depends on three operative components:

- *Multimedia Records (content)*: the *documental layer* of resources evocative of experiences shared by the various participants, consisting of a structured database of the records and their metadata which allow correlated forms of presentation.
- *Collective Memories (visualizations)*: the *aggregation and correlation layer*, which automatically organizes the records based on defined filters and semantic connections. The resulting visualizations are dynamic and customizable, working as triggers for remembrance and content discovery.
- *Digital Community (users)*: the *social and emotional layer* that feeds the system through social engagement and content sharing. This digital community includes different types of users whose motivation and level of participation is variable (from institutions and event promoters that disseminate content and foster its activation through mediation activities, to more passive and contemplative users, as opposed to others who are more active in contributing with their records).

The participatory platform is the central element of the model, whose digital mediation enables the visualization and co-creation of collective memories, based on shared records. This social dynamic of content sharing and enrichment leads to the constitution of *digital archives* and the consolidation of *online communities*, around common interests, and experiences, with the possibility of replication and adaptation to different contexts. The identified use contexts are cultural content, places, or events dynamized by communities, local organizations, or municipality facilities such as museums, archives, libraries, natural parks, etc.

The model's components articulate according to three platform principles (Figure 2) which are mutually interdependent – mediation, visualization, and participation. The visualization and participation mechanisms only exist due to the mediation mechanisms provided by the digital platform, and this mediation is focused on the operationalization of the other two mechanisms. Likewise, visualization mechanisms are inapplicable if they are not activated by participation mechanisms that feed the system with records and through social dynamics that enrich and contextualize it.

Mediation Principle
The platform should support activities that encourage remembrance and social dynamics around the digital archive of experiences' records shared by various users, to better represent these collective memories and consolidate the community around them
Visualization Principle
The platform should offer automatic and customizable ways of aggregating, organizing and presenting shared resources to enhance their critical and creative exploration from multiple perspectives
Principle of Participation
The platform should provide functional features for contribution, exploration, interaction, and creation, to enhance social dynamics of enrichment, remix, and co-creation, based on the content shared by multiple users

Fig. 2. Principles of the MNEMOSCOPE model: mediation, visualization, and participation.

The three principles unfold into detailed guidelines, providing good practices for the sustainability of digitally mediated participatory projects and the development of prototypes applied to specific contexts. Conceptually, the model was designed in a flexible way to be able to contemplate two approaches: 1) to be one broad scope platform, in which several 'memory projects' can be created and assigned to different themes/events with targeted audiences; 2) to allow the creation of different platforms for specific projects/events, designed from scratch to address specialized needs and purposes, based on the principles underlying the model. Such adaptations to niche contexts and uses are welcome, as they may contribute to improving and expanding the model, as an increment, extension, or derivation, from a perspective of open access to knowledge and scientific cooperation.

3 Final considerations and future work

Through the intersection of different methods and instruments (study of multiple cases; prototyping; and interviews with experts), it was possible to design a

technology-mediated model for co-creation and visualization of collective memories, guided by principles of mediation, visualization, and participation. The model is based on a digital platform which constitutes simultaneously a collaborative archive and a digital community around records of shared events and experiences. Compared to other participatory projects and platforms (including social media), the differentiating aspect focuses on the visual and semantic correlation of multimedia records to be socially enriched, instigated by correlated visualizations that depict different points of view and exploratory possibilities.

In this regard, the core of the digital platform stands out as an instrument of storytelling based on the sharing of visual records and oral testimonies, namely mobile capture during events and their real-time integration in a structured database. Also, it is an exploratory instrument based on a semantic and visual correlation system, which automatically generates customizable visualizations to highlight connections between content shared by various users. The model's versatility also contemplates its application for research purposes regarding visual studies, participatory culture studies and contemporary archiving practices, as a complementary tool for documenting and analysing participation dynamics and shared multimedia content.

The potential that emerged from this research, which culminated in the model, was translated into an R&D project approved for funding. The POLARISCOPE project was initiated in 2023 and will run over a period of three years, focused on the implementation and user testing of a functional platform in field trials during cultural events and experiences in Portugal. The informal and institutional usages of the platform in different territories and events will allow the co-creation and transfer of innovation and knowledge to the community, to support meaningful participation and instigate critical debate and creativity around collective memories.

Acknowledgements

The research is funded by FCT - Fundação para a Ciência e a Tecnologia (Grant nr. SFRH/BD/132780/2017). The authors acknowledge the collaboration of the experts.

References

1. Jenkins, Henry: *Convergence Culture: Where Old and New Media Collide*. New York University Press, New York (2008).
2. Alexander, B. (2011). *The New Digital Storytelling - Creating Narratives with New Media*. Santa Barbara, California: Praeger.
3. Segel, E., Heer, J.: Narrative Visualization: Telling Stories with Data. In: *Visualization and Computer Graphics*, IEEE Transactions on 16.6., pp. 1139–1148 (2010).
4. Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages (VL '96)*, pp. 336–343. USA: IEEE Computer Society (1996).
5. Manovich, L.: *The Language of New Media*. The MIT Press, Cambridge, Massachusetts (2001).

6. Cairo, A.: *The Functional Art – an introduction to information graphics and visualization*. New Riders, USA (2013).
7. Huybrechts, L. (Ed.): *Participation is Risky – Approaches to Joint Creative Processes*. Valiz, Amsterdam (2014).
8. Drucker, J.: *Graphesis: Visual knowledge production and representation* (2011), http://peterahall.com/mapping/Drucker_graphesis_2011.pdf, last accessed 2023/02/10.
9. Dijck, J. V.: *Mediated memories in the digital age*. Stanford University Press, Stanford (2007).
10. Ernst, W.; *Digital Memory and the Archive*. University of Minnesota Press, Minneapolis, London (2013).
11. Ienken, I., Neilson, T., Rheams, D.: *Research Methods for the Digital Humanities*. Palgrave Macmillan (2018).
12. Manovich, L.: *Cultural Analytics*. The MIT Press, Cambridge, London (2020).
13. Pousman, Z., Stasko, J., Mateas, M.: *Casual Information Visualization: Depictions of Data in Everyday Life*. In: *Visualization and Computer Graphics*, IEEE Transactions on 13.6, pp. 1145–1152 (2007).
14. Whitelaw, M.: *Generous Interfaces for Digital Cultural Collections*. In: *Digital Humanities Quarterly* 9,1 (2015).
15. Surowiecki, J.: *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday, New York (2004).
16. Brabham, D. C.: *Crowdsourcing*. The MIT Press, Cambridge, London (2013).
17. Velhinho, A.: *Sobre a Influência da Visualização e da Participação na Cultura Visual em Rede do Século XXI e na Cocriação Digital de Memórias Coletivas* [Doctoral thesis]. Faculdade de Belas-Artes da Universidade de Lisboa (2023).
18. Glaser, B. G., & Strauss, A. L.: *The discovery of Grounded Theory - strategies for qualitative research*. Routledge, London, New York (2017).
19. Velhinho, A., Almeida, P.: (2021). *Sharing and visualizing collective memories – contexts and strategies for a participatory platform*. In: Abásolo M., Abreu J., Almeida P., Silva T. (eds) *Communications in Computer and Information Science*, pp. 3-14. Springer International Publishing (2020).
20. Silva, A. M.: *Modelos e Modelizações em Ciência da Informação: O Modelo eLit.pt e a investigação em literacia informacional*. *Prisma.com* (Portugal), n. 13, pp. 298–353. (2010).
21. Sampaio, C. P., Martins, S. B.: *Projetos de pesquisa e desenvolvimento em design, sustentabilidade e inovação: bases teóricas para a contribuição do design*. *Design, Artefatos e Sistema Sustentável*, pp. 35–58, (2018).
22. Sousa, F.: *The Participation in the Safeguarding of the Intangible Cultural Heritage. The role of Communities, Groups and Individuals*. *Memória Imaterial CRL*, Alenquer (2018).

Conformación dinámica de visitas virtuales con vista panorámica

Carlos-Alberto Flores-Sanjurjo, Joaquín-Danilo Pina-Amargós^[0000-0003-4619-849X] and Ariel-Alfonso Fernández-Santana

Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE)
jpina@ceis.cujae.edu.cu

Resumen Las visitas virtuales permiten simular la presencia de los seres humanos en lugares determinados desde la comodidad del hogar, a través de la inmersión. En la actualidad son utilizadas por muchas personas e incluso por grandes empresas. Existen múltiples softwares que permiten la creación de estas visitas, pero en su mayoría sólo permiten la visualización de escenas estáticas cuyo contenido no es posible modificar, además que son softwares privativos y no es posible mejorarlo o incorporar nuevas funcionalidades o su uso está prohibido para algunos países con sanciones unilaterales impuestas por potencias extranjeras. Para resolver ese problema se desarrolló una investigación en aras de seleccionar una tecnología que permita desarrollar un software que brinde a Cuba independencia tecnológica y cumpla con los requisitos necesarios. De acuerdo a esto, se decidió utilizar las tecnologías libres asociadas a JavaScript, empleando Node, Nest.js, Mongo DB, en el desarrollo de un backend para la persistencia de los datos y para las funcionalidades visuales asociadas a la realidad virtual se escogió three.js, biblioteca libre y de código abierto de JavaScript que cuenta con una gran cantidad de funciones disponibles. Como la solución demanda escasos recursos, los resultados pueden ser aplicados a la televisión digital interactiva en cajas decodificadoras inteligentes de bajas prestaciones para lograr extender su utilización en la educación y cultura de los pueblos. Este trabajo contribuye al desarrollo de diferentes sectores de la sociedad y fomenta el crecimiento de la soberanía tecnológica de los países.

Palabras claves: realidad virtual, visita virtual, software libre.

1. Introducción

Desde el surgimiento de la informática, el ser humano ha soñado con la posibilidad de realizar una inmersión en mundos alternativos creados de forma virtual, para experimentar sensaciones que en la vida real no podría sentir. La realidad virtual encierra una esfera de conocimiento que es la encargada de la representación de escenas, multimedia y objetos, con una estrecha relación entre sí, para conformar una experiencia real. Diversos autores definen a la realidad virtual como la presencia del humano en el espacio generado por el ordenador de manera interactiva, donde el usuario entra en contacto junto al ordenador con un mundo prácticamente legítimo [8].

En la actualidad, la presencia de la realidad virtual es algo común. Por ejemplo, se puede encontrar en diseños arquitectónicos, en videojuegos, en simuladores de aviones y automóviles, en simuladores de tiro, en anuncios realizados por las empresas publicitarias y en los sectores de la educación, el arte y el entretenimiento [6]. También se pueden encontrar sus aplicaciones en la medicina como en los procedimientos quirúrgicos de mínimo acceso, en la anatomía y en la rehabilitación [5].

Una de las formas de representar un entorno a través de la realidad virtual son las visitas virtuales, método de representación de la realidad que se ha vuelto muy popular en los últimos años [9]. Se usan en la rama inmobiliaria para promover sus productos sin necesidad de la presencia física de los clientes ya que se pueden difundir a través de internet [4]. También se usan en el ámbito educativo y artístico a través de la representación de museos y centros de interés, para enaltecer el patrimonio cultural de la nación. Además, en la criminalística brindan la capacidad de reconstruir una escena del crimen para esclarecer los sucesos [4].

Existen aplicaciones para la creación de visitas virtuales, las cuales de manera interactiva representan un lugar determinado mediante imágenes panorámicas o modelos 3D [10]. Además, permiten la interacción de los usuarios con el entorno para navegar entre escenas o visualizar determinados puntos de interés. Otra de las características que poseen es la navegación a través de una ruta automática por la localización [9]. Estas herramientas brindan la capacidad de incorporar a los recorridos audios, imágenes, videos, textos e indicaciones. Con el objetivo de representar un mundo virtual más real, las visitas virtuales utilizan tres aspectos fundamentales de la realidad virtual: la inmersión, el tiempo real y la interacción [3].

En la actualidad, la web es el medio más utilizado por los usuarios a través de cualquier dispositivo que cuente con la capacidad de conectarse a internet, logrando el acceso a cualquier parte del mundo en cuestión de segundos. En ella podemos encontrar todo tipo de información que permite la superación y, además, facilita la realización de cualquier tipo de investigación. Por estas razones, han surgido tecnologías innovadoras para mejorar cada día las experiencias de los usuarios en la web.

En el presente trabajo se expone el desarrollo de un sistema que permite la visualización y la conformación dinámica de visitas virtuales con vistas panorámicas, utilizando tecnologías basadas en software libre en un entorno web. A continuación se abordan los antecedentes encontrados en la literatura especializada y se realiza una fundamentación de la selección tecnológica escogida. Luego se expone la solución propuesta para demostrar la viabilidad de la conformación dinámica y visualización de visitas virtuales en la web, utilizando tecnología libre, a través de un prototipo funcional. Para finalizar se exponen y discuten los resultados alcanzados con la presente investigación.

Los beneficios de este trabajo se reflejarán en el ámbito económico, su aplicación en el turismo permitirá que los hoteles del país sean promovidos, lo cual puede traer como resultado un mayor número de clientes, constituyendo una mayor

fuentes de ingresos. Además, con su realización se estará adquiriendo soberanía tecnológica, por lo que utilizando el software libre que aborda la investigación se podrán desarrollar aplicaciones similares a estas, sin depender de terceros, y, sobre todo, sin comprar en el mercado internacional. El empleo frecuente de visitas virtuales en el país resultará beneficioso para enaltecer el patrimonio cultural e histórico de la nación, promoviendo lugares históricos, museos, galerías de arte, centros turísticos, zonas silvestres, entre otros. También, constituyen un modo de incentivar la educación de los jóvenes, acercándolos a la historia de Cuba de una manera visual, moderna y atractiva.

2. Antecedentes y selección tecnológica

Se realizó un estudio de múltiples aplicaciones libres que pudieran cumplir los parámetros abordados en esta investigación. A continuación se muestran junto con sus principales limitaciones:

1. Marzipano [1]
 - No permite la persistencia de datos de las visitas virtuales.
 - Solo es compatible con los navegadores Firefox y Chrome.
 - No permite el uso de multimedia cargada desde la web.
2. Pano2VR [2]
 - Tiene una demo gratuita con muy pocas funcionalidades.
 - Para acceder a todas sus funcionalidades se debe adquirir una licencia de pago.

Por tanto, las aplicaciones libres no cuentan con todas las funcionalidades necesarias para el desarrollo y visualización de las visitas virtuales interactivas, ni permiten su conformación dinámica. Por su parte, otras aplicaciones brindan una mayor cantidad de funcionalidades pero no son accesibles para Cuba o son privativas.

La investigación propone un método de construcción dinámica para las visitas donde se pueda modificar el comportamiento de todos los puntos de interés en la misma. Para evitar que la aplicación cliente se sobrecargue, todos los datos y multimedia serán almacenados en un servidor web al cual se accederá mediante la red. Esto permitirá a su vez que cualquier usuario pueda agregar o modificar información en la misma aplicación y se actualizará de manera automática.

Teniendo en cuenta los antecedentes planteados, este trabajo propone la adopción de tecnologías y estándares abiertos en la Televisión Digital Interactiva para la educación y el entretenimiento. Estas tecnologías fueron probadas en dos prototipos de software, una multimedia y un videojuego, donde se obtuvieron buenos resultados que fueron bien valorados por los usuarios encuestados.

Se desea que el sistema sea asequible desde distintos dispositivos por lo que se decidió desarrollar la misma en entorno web, permitiendo de esta forma su visualización utilizando un navegador tanto desde un ordenador como desde un dispositivo móvil o *tablet*.

Para ilustrar la selección tecnológica realizada se presenta la Tabla 1 donde se muestra una comparación de las tecnologías web 360 referidas anteriormente en los antecedentes de la investigación.

Tabla 1: Tecnologías web 360

Atributo	<i>A-frame</i>	<i>Pannellum</i>	<i>Marzipano.js</i>	<i>Three.js</i>
Licencia	<i>MIT</i>	<i>MIT</i>	<i>MIT</i>	<i>MIT</i>
Curva de aprendizaje	Alta	Media	Baja	Media
Documentación	Media	Media	Media	Alta
Última Versión Estable	04/02/2021	26/11/2019	14/11/2021	30/04/2022

Tras realizar el análisis anterior se decidió utilizar *Three.js* para el manejo de la realidad virtual en el lado del cliente. Los criterios en los que se basó esta decisión se presentan a continuación:

1. Presenta licencia *MIT*.
2. A pesar de tener una curva de aprendizaje media el autor ya había trabajado con anterioridad con esta biblioteca, por lo que se encontraba muy familiarizado con ésta.
3. Presenta mucha documentación, tutoriales y guías para desarrollar aplicaciones.
4. La última versión estable fue liberada en abril del 2022.

En la Tabla 2 se muestra una comparación de las tecnologías para el desarrollo del lado del cliente referidas en el epígrafe anterior.

Tabla 2: Tecnologías *frontend*

Atributo	<i>AngularJS</i>	<i>VueJS</i>	<i>React.js</i>
DOM	<i>Regular DOM</i>	<i>Virtual DOM</i>	<i>Virtual DOM</i>
Curva de aprendizaje	Alta	Baja	Baja
<i>Packaging</i>	Débil	Fuerte	Fuerte
<i>MVC</i>	Sí	Inspirado	Solo capa de vista

Tras realizar el análisis anterior se decidió utilizar *React.js* para el desarrollo del lado del cliente. Los criterios en los que se basó esta decisión se presentan a continuación:

1. Utiliza un *DOM* virtual que permite un gran rendimiento.
2. Presenta una curva de aprendizaje baja y utiliza en gran medida las funcionalidades de *JavaScript*.
3. No utiliza el patrón *MVC*, sin embargo, dada la arquitectura a utilizar que se analizará en el siguiente capítulo, esto no representa una desventaja.

En la Tabla 3 se muestra una comparación de los *frameworks* para el desarrollo del lado del servidor referidos en el epígrafe anterior.

Tabla 3: Tecnologías *backend*

Atributo	<i>Asp.Net</i>	<i>Spring</i>	<i>NestJS</i>
Arquitectura	Sí	Sí	Sí
Utilización de recursos	Alto	Alto	Bajo
Cantidad de código	Media	Alta	Baja
Lenguaje de programación	<i>C Sharp</i>	<i>Java</i>	<i>JavaScript</i>

Tras realizar el análisis anterior se decidió utilizar *NestJS* para el desarrollo de la aplicación en el lado del servidor. Los criterios en los que se basó esta decisión se presentan a continuación:

1. Brinda una arquitectura predeterminada.
2. El factor que más influyó en la elección es que consume muy pocos recursos del sistema.
3. Permite desarrollar aplicaciones utilizando poca cantidad de código, lo cual facilita la realización de pruebas.
4. Utiliza el lenguaje de programación *JavaScript*, lo cual implica que se puede realizar todo el sistema en un solo lenguaje.

3. Solución propuesta

Se necesita demostrar la viabilidad de la conformación dinámica y visualización de visitas virtuales en la web, utilizando tecnología libre, a través de un prototipo funcional. Se debe incluir la posibilidad de interactuar con el entorno, navegar entre las escenas, visualizar imágenes y mostrar información sobre elementos de interés presentes en la escena.

Al no identificarse procesos claramente definidos en la modelación del negocio, se consideró la realización de una descripción de las entidades y relaciones involucradas en este. Por tal razón, no existe diagrama de caso de uso del negocio y se tomó la decisión de utilizar un modelo de dominio para la representación. En la Figura 1 se muestra la representación gráfica del dominio del sistema junto a un glosario de términos en la Tabla 4 para una mejor comprensión del sistema.

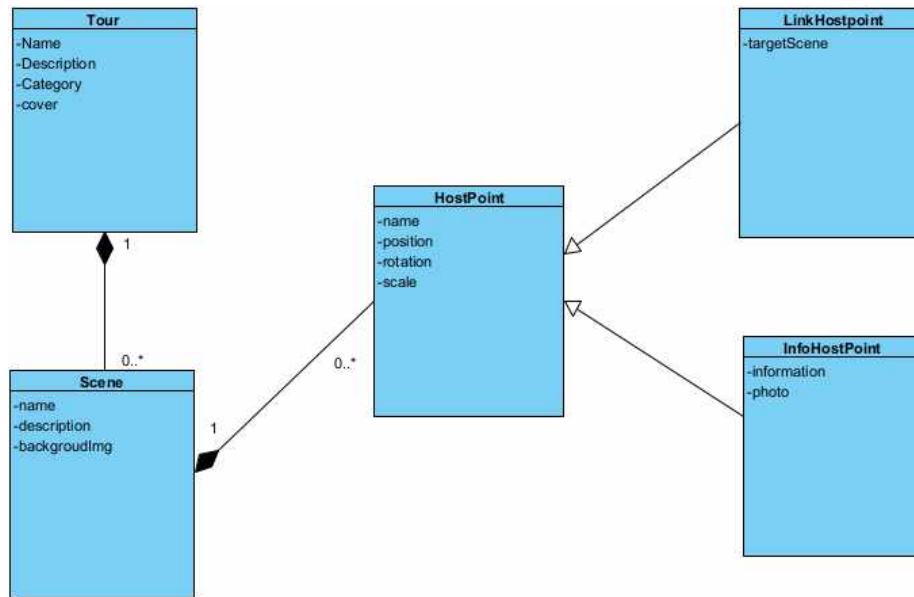


Figura 1: Diagrama de dominio.

Tabla 4: Definición de entidades y conceptos principales

Entidad	Descripción
<i>Tour</i>	Representa un proyecto de visita virtual, posee el conjunto de escenas que conforman toda la visita.
<i>Scene</i>	Representa una escena, es el conjunto de todos los diferentes tipos de elementos que conforman la visita virtual.
<i>LinkHostpoint</i>	Es representado como un componente sphere en Three.js y permite el desplazamiento entre las escenas del tour.
<i>InfoHostpoint</i>	Es representado como un componente sphere en Three.js y permite mostrar una información determinada al usuario.

En la tabla 5 se muestran los requisitos funcionales del sistema.

Tabla 5: Requisitos funcionales.

Código	Nombre	Descripción
C-01	-Visualizar tour virtuales	El visitante es capaz de visualizar las visitas virtuales que ha creado con anterioridad.
C-02	-Conformar tour virtuales	El editor es capaz de crear una visita virtual, al tocar el botón Crear Visita Virtual, se le mostrará un formulario, en el que deberá escribir nombre, descripción, categoría de la visita y, además, debe seleccionar una foto para la portada de la visita y así comenzar a crear el tour.
C-03	-Cargar escena	Al acceder a un tour virtual el sistema automáticamente carga las escenas asociadas a este tour.
C-04	-Cargar <i>infoHostPots</i>	Al acceder a una escena el sistema automáticamente carga los <i>infoHostPots</i> asociados a la escena actual.
C-05	-Cargar <i>linkHostPots</i>	Al acceder a una escena el sistema automáticamente carga los <i>linkHostPots</i> asociados a la escena actual.
C-06	-Gestionar escena	El editor es capaz de crear, editar y eliminar las escenas dentro del tour virtual, al pulsar en los botones respectivos a cada acción.
C-07	-Gestionar <i>linkHostPots</i>	-El editor es capaz de crear, editar y eliminar los <i>linkHostPots</i> dentro de la escena al pulsar en los botones respectivos a cada acción.
C-08	-Gestionar <i>infoHostPots</i>	-El editor es capaz de crear, editar y eliminar los <i>infoHostPots</i> dentro de la escena al pulsar en los botones respectivos a cada acción.
C-9	-Registrar usuarios	-Para acceder al sistema todos los usuarios deben registrarse con anterioridad.
Continúa en la página siguiente...		

Tabla 5 – Continuación de la página anterior.

Código	Nombre	Descripción
C-10	-Autenticar <i>linkHostPot</i>	- Para acceder al sistema el usuario debe proporcionar un nombre de usuario y una contraseña que fueron registradas anteriormente.

Como se puede apreciar en la Tabla 6 uno de los requisitos de calidad del sistema es la disponibilidad. Para garantizar el cumplimiento de este requisito se decidió desarrollar la aplicación utilizando una arquitectura basada en micro-servicios.

Los microservicios se refieren a una arquitectura orientada a servicios que busca descomponer una aplicación en diferentes servicios, con el objetivo de obtener alta disponibilidad, bajo acoplamiento, descentralización y tolerancia a fallas. Es una arquitectura ideal para escenarios de alto tráfico y alta disponibilidad.

Tabla 6: Requisitos de calidad

Código	Nombre	Descripción
C-01	Usabilidad	Se debe garantizar una interfaz simple e intuitiva.
C-02	Navegabilidad	Se debe garantizar URLs legibles y fácil de recordar, así como poder navegar entre página sin usar las opciones del navegador.
C-03	Tolerancia a fallas	Se debe garantizar que los errores y eventos inesperados no impliquen el colapso de todo el sistema
C-04	Alta disponibilidad	Se debe garantizar el funcionamiento del sistema en todo momento del día.

En la Figura 2 se muestra el diagrama de casos de uso del sistema el cual contiene las funcionalidades a implementar que debe brindar el sistema.

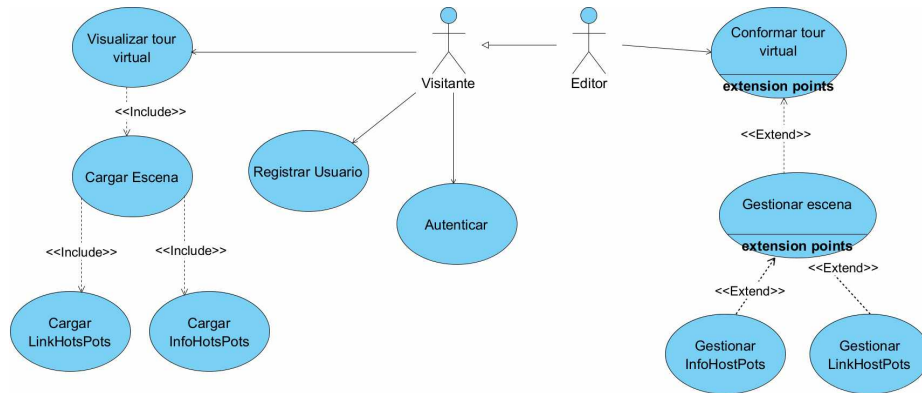


Figura 2: Diagrama de caso de uso del sistema.

Para la implementación de esta arquitectura en esta solución, se siguió una estrategia de diseño basada en núcleo de negocio. Esto significa que cada entidad identificada en el modelo del dominio, representa un microservicio, como se puede apreciar en la estructura de carpetas mostrada en la figura 3.

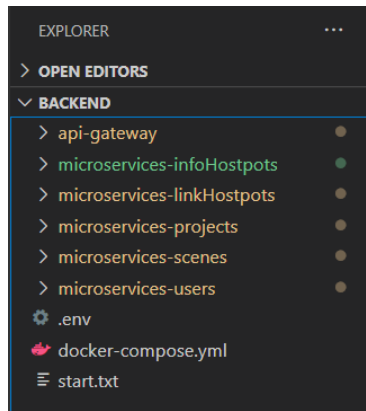


Figura 3: Estructura de carpetas del *backend*.

Cada microservicio puede estar construido en un lenguaje diferente, lo que facilita el trabajo en equipo. En el siguiente epígrafe se demuestra el cumplimiento de los patrones de diseño que garantizan una mejor calidad en la construcción de esta arquitectura.

Principios de diseño

1. *Open – Close*: Este principio se pone de manifiesto en la mayor parte de la solución, pues todos los patrones de diseño de la propuesta anterior cumplen con él.
2. *Once and Only Once Rule*: Este principio se puede evidenciar en la creación de los *InfoHostpots*, procedimiento que se implementa siguiendo el patrón *Template Methods*.
3. Diseñar hacia las interfaces, no hacia las implementaciones: Al igual que el *Open- Close* este principio se evidencia en todos los patrones de diseño empleados en la propuesta de solución.
4. *Hollywood*: Este es uno de los principios que se ponen de manifiesto al usar en la solución el patrón *Template Methods*.
5. Encapsular la variabilidad: Principio que se pone de manifiesto al emplear el patrón *Observer* y el patrón *Prototype* en la propuesta de solución.
6. Ley de *Deméter*: Al igual que el principio anterior, su cumplimiento se manifiesta en el uso de los patrones *Observer* y *Prototype*.
7. Bajo acoplamiento: Igualmente, es otro de los principios que se evidencian por haber empleado en la propuesta de solución los patrones *Observer* y *Prototype*.

Para dar cumplimiento al objetivo de esta investigación se decidió desplegar la aplicación en un entorno web. El cual está dirigido a usuarios que accedan a un servidor web donde estará alojado el cliente de la visita virtual a través de un navegador, éste a su vez, se comunicará con los servidores donde estarán alojados los diferentes microservicios del *backend* y sus respectivas bases de datos a través del protocolo *HTTP*.

En la Figura 4 se muestra el diagrama de despliegue del sistema como una aplicación web basada en microservicios.

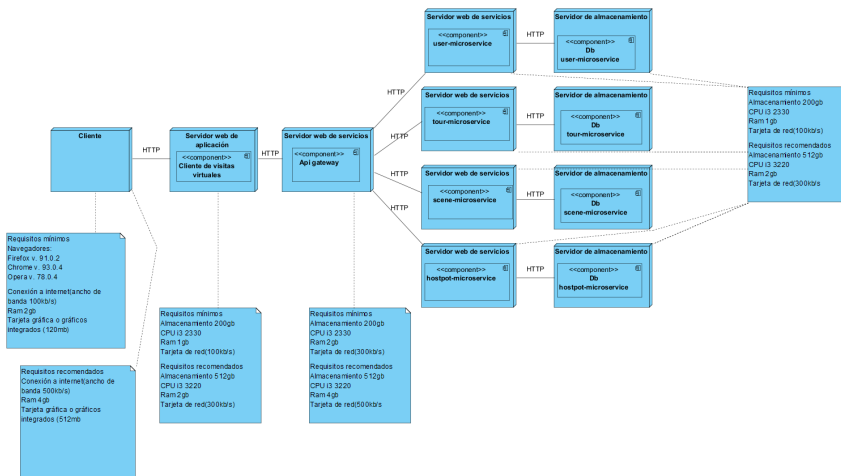


Figura 4: Diagrama de despliegue.

La viabilidad de la conformación dinámica de visitas virtuales interactivas en entornos de desarrollo libre es irrefutable, tanto práctico como teórico y metodológico. Al demostrar que se poseen los medios para lograrlo se estaría ganando en independencia tecnológica, ofreciendo una opción que, de otro modo, resultaría muy costosa.

Según [7], [4] el uso de las visitas virtuales como método de aprendizaje motiva a los estudiantes, además, de demostrarse que es mayor el conocimiento que se afianza. También, se debe tener en cuenta que la digitalización del patrimonio de la nación es importante y las visitas virtuales a museos, monumentos y similares contribuyen a preservarlo volviéndolo a la vez asequible a una mayor audiencia de una forma más atractiva y moderna.

Basándose en lo anterior, se considera que la realización de este trabajo es viable, debido a que las tecnologías seleccionadas permiten cumplir con el objetivo general planteado. Las necesidades básicas serían contar con un conjunto de imágenes panorámicas y un dispositivo que permita la ejecución de un navegador web para poder acceder a la aplicación.

4. Análisis de resultados

Una vez culminado el proceso de prueba se procede a analizar los resultados obtenidos:

- El prototipo diseñado abarcó todos los requisitos funcionales y fue un elemento clave para la ejecución de los casos de prueba.
- Las pruebas asociadas a los casos de uso relativos al procesamiento de las interacciones de los *infoHostpot* y *linkHostpot* arrojaron resultados exitosos. Esto evidencia una correcta implementación de las funcionalidades de interacción en la visita virtual.
- Las pruebas asociadas a los casos de uso relacionados a la creación de las visitas virtuales y sus componentes, demostraron que se cumple la conformación dinámica de los recorridos virtuales.
- Las pruebas realizadas al caso de uso asociado a la visualización de las visitas virtuales resultaron exitosas, demostrando la correcta implementación de las funcionalidades de visualización y navegación libre a través de las escenas.

En la Tabla 7 se muestran las características del ordenador utilizado para realizar las pruebas.

Flujo de trabajo Al comenzar a trabajar con el sistema el usuario deberá autenticarse, luego de esto, se visualiza en la pantalla todos los recorridos virtuales que ha creado dicho usuario. En este punto se puede elegir si crear una nueva visita virtual o acceder a alguna que haya sido creada previamente. En la Figura 5 se puede apreciar lo anterior.

Tabla 7: Características de la computadora en la que se realizaron las pruebas

Recursos	Prestaciones
CPU	Intel(R) Core(TM) i3-8100 CPU 3.60GHz
Motherboard	H310M H 2.0
GPU	Intel UHD Graphics 630
RAM	4 GB DDR4 1200 MHz
HDD	1 TB

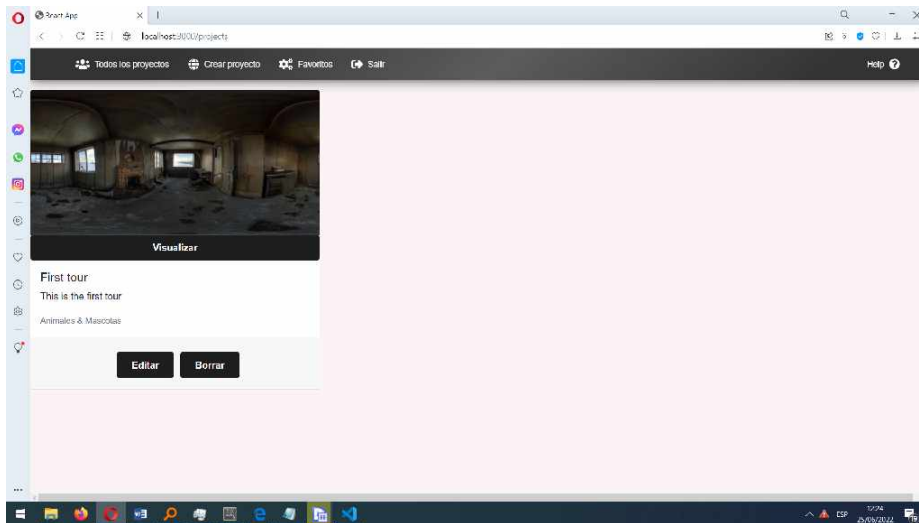


Figura 5: Pantalla del sistema *Lista de visitas virtuales*.

Al acceder a *editar el tour*, se muestra la pantalla que va a permitir conformar la visita al usuario. Es necesario familiarizarse con esta vista antes de empezar a editar el recorrido. En la ventana que se muestra en la Figura 6 y en la Figura 7 se puede apreciar que:

- En la barra inferior se encuentran todas las escenas creadas.
- En la barra lateral izquierda hay un inspector que muestra dinámicamente los datos de los componentes que intervienen en la escena.
- En la barra superior hay un botón *Elementos* que permite la creación de los puntos de interés en la escena.
- En el resto de la pantalla se encuentra la escena con su imagen panorámica correspondiente.
- En la pantalla aparece una ventana con el formulario para crear una escena al pulsar el botón añadir.

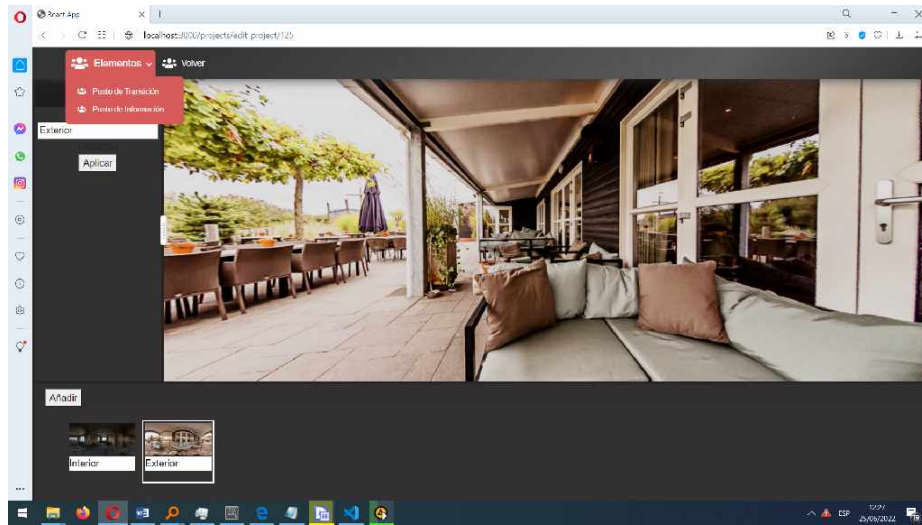


Figura 6: Pantalla del sistema *Editor de visitas virtuales*.

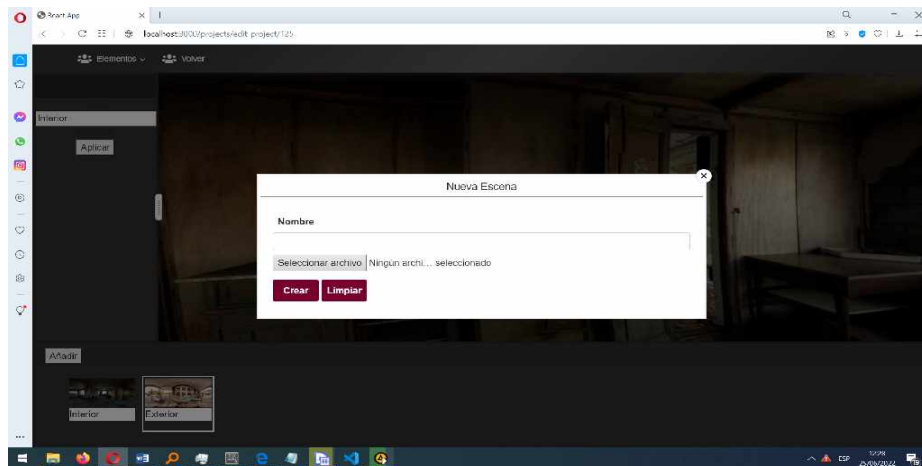


Figura 7: Pantalla del sistema *Editor de visitas virtuales*.

Como se puede apreciar el usuario es capaz de crear una visita virtual sin necesidad de codificar. Dichas visitas pueden ser personalizadas utilizando cualquier foto panorámica que se desee. Además, la inclusión de puntos de información permiten una experiencia interactiva, debido a que se les puede incorporar la información que desee el usuario. De la misma forma, al añadir los puntos

de transición se posibilita la capacidad de realizar un recorrido a varias escenas dentro de la misma visita.

Luego de finalizada la conformación dinámica de la visita, el usuario es capaz de visualizarla al pulsar en el botón visualizar que se encuentra en la pantalla de *Lista de visitas virtuales*. Ésta vista propicia una inmersión parcial en el recorrido, permitiendo al usuario interactuar con los elementos del entorno creados con anterioridad. En la Figura 8 se puede apreciar dicho planteamiento.

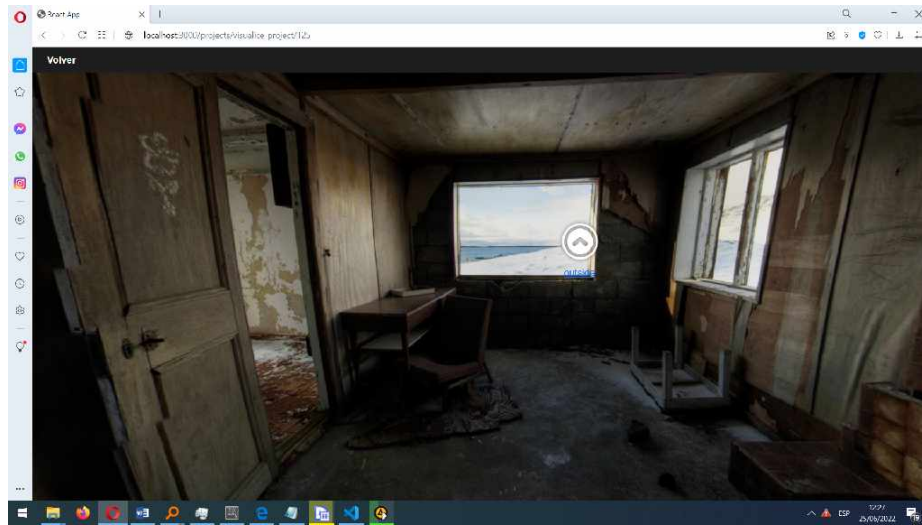


Figura 8: Pantalla del sistema *Visualización de visitas virtuales*.

5. Conclusiones

Mediante la realización de esta investigación se cumplieron los objetivos trazados arribando a las siguientes conclusiones: El estudio del estado del arte demostró la necesidad de contar con un sistema capaz de conformar y visualizar una visita virtual, así como de interactuar con elementos dentro de ésta. Las visitas virtuales tienen un amplio uso, destacando en el marketing y la industria turística, además, son un buen medio de aprendizaje y tributa a preservar y enaltecer el patrimonio cultural e histórico de la nación. Las características principales con las que debe contar una visita virtual 360 son: navegar libremente entre las escenas y mostrar información sobre puntos de interés en forma de texto y elementos multimedia. El desarrollo de visitas virtuales 360 interactivas basadas en las tecnologías de *JavaScript* en un entorno libre es factible, ofreciendo soberanía tecnológica al país. Se logró desarrollar la conformación dinámica de la visita virtual al permitir la modificación de todos sus componentes sin modificar el código. El uso de las pruebas con enfoque de caja negra permitió comprobar

que las funcionalidades que demuestran la visualización y conformación dinámica se cumplieron satisfactoriamente, garantizando la calidad del producto. El sistema implementado satisface la problemática planteada al inicio del trabajo en un entorno de desarrollo libre.

Para darle continuidad a este trabajo se recomienda: Permitir exportar cada visita creada como un proyecto de *react* para su incorporación en otros sistemas web. Implementar una aplicación de escritorio o móvil para permitir el trabajo sin conexión del sistema. Permitir que los usuarios puedan compartir sus visitas a través del mismo sistema o utilizando las redes sociales.

6. Agradecimientos

Esta investigación ha sido apoyada por el Fondo Fiduciario Pérez-Guerrero para la Cooperación Sur-Sur (PGTF) del Programa de las Naciones Unidas para el Desarrollo (UNDP) proyecto INT/19/K08 y el Ministerio de Ciencia, Tecnología y Medio Ambiente de Cuba (CITMA) proyecto NPN223LH006-005 de TVDi.

Referencias

1. <https://www.marzipano.net> (2014), accedido 10-5-2022
2. <https://ggnome.com/pano2vr> (2014), accedido 10-5-2022
3. Burdea, G.C., Coiffet, P.: Virtual reality technology. *International Journal of e-Collaboration* **2**(1), 61–64 (2006)
4. Chiao, H.M., Chen, Y.L., Huang, W.H.: Examining the usability of an online virtual tour-guiding platform for cultural tourism education. *Journal of Hospitality, Leisure, Sport & Tourism Education* **23**, 29–38 (2018)
5. Dimbwadyo-Terrer, I.: Realidad virtual en procedimientos quirúrgicos: Nuevos entornos de aplicación. *NeuroRehabNews* (Octubre) (2017)
6. Kim, J.S., Kim, S.J., Park, S.N., Shin, J.W., Kwon, S.Y.: A study on the development of vr-based education and culture program in public libraries. *Journal of the Korean Society for information Management* **38**(2), 87–112 (2021)
7. Loures Brandão, G.V., Henriques do Amaral, W.D., Rabite de Almeida, C.A., Barroso Castañon, J.A.: Virtual reality as a tool for teaching architecture. In: *International Conference of Design, User Experience, and Usability*. pp. 73–82. Springer (2018)
8. Mazuryk, T., Gervautz, M.: *Virtual reality-history, applications, technology and future* (1996)
9. Pincay Bermello, V.R., Rivas Rodríguez, M.J.: Tour virtual interactivo 360 de las instalaciones de la Universidad de Guayaquil. B.S. thesis, Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas ... (2021)
10. Widiyaningtyas, T., Prasetya, D.D., Wibawa, A.P.: Web-based campus virtual tour application using orb image stitching. In: *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. pp. 46–49. IEEE (2018)

Audiovisual Consumption |

GoWatch: una aplicación de televisión interactiva para seguir el consumo audiovisual en diferentes plataformas de streaming

Antoni Oliver^[0000-0002-2495-5245], Antoni Bibiloni^[0000-0002-7359-8280] and Andrea Morey.

Laboratori de Tecnologies de la Informació Multimèdia (LTIM), Departament de Ciències Matemàtiques i Informàtica, Universitat de les Illes Balears, Palma, España
antoni.oliver@uib.es, toni.bibiloni@uib.es,
andrea.morey1@estudiant.uib.es

Abstract. Siguiendo la popularidad de Netflix, muchas plataformas de streaming han aparecido recientemente. Hoy en día, hay una gran cantidad de contenido multimedia, series de televisión y películas que están disponibles en diferentes sitios de streaming. Esto introduce dos problemas: primero, es difícil saber en qué sitio se puede consumir un contenido; y segundo, es difícil saber qué contenido se ha consumido y, en el caso de series de televisión, hasta qué episodio. En este artículo presentamos GoWatch, una aplicación de televisión interactiva que intenta ayudar a sus usuarios en estas dos áreas siguiendo a qué plataformas de streaming están suscritos sus usuarios y qué contenidos consumen. GoWatch también sirve como un catálogo público interactivo del contenido multimedia disponible en estas plataformas; introduce características sociales con amigos y recomendaciones; y notifica a sus usuarios cuando contenido adicional está disponible (por ejemplo: cuando la siguiente temporada de una serie de televisión ha sido publicada). Discutimos el diseño, implementación y resultados del Desarrollo de esta aplicación y concluimos con algunas reflexiones del trabajo realizado.

Keywords: plataformas de streaming, televisión interactiva, Samsung Smart TV, Tizen.

1 Introducción

A principios de la década de 2010, los televisores inteligentes o *Smart TV* llegaron al mercado de los televisores para revolucionarlo por completo. En 2021, aproximadamente el 60% de la población española ya disponía de un televisor inteligente en sus hogares [1].

La clave del éxito de esta nueva generación de televisores es que además de incorporar las funcionalidades básicas de un televisor tradicional, tienen integrados Internet y características de la web 2.0. Gracias a estas nuevas incorporaciones, los televisores inteligentes son capaces de proporcionar muchas más formas de entretenimiento. Un claro ejemplo de ello son las plataformas *streaming*, que permiten la transmisión de

contenido multimedia a través de Internet sin necesidad de los proveedores tradicionales.

En 2015, aterrizó en España Netflix, una de las plataformas *streaming* más populares a nivel mundial y desde aquel entonces la seguirían muchas más. A partir de este momento, empezó a producirse un cambio inevitable en la manera de consumir televisión, llegando al punto en el que, tal y como se constata en un artículo del diario digital El Independiente [2], “el consumo televisivo había sufrido en 2021 una bajada histórica (...). No se registraban datos tan bajos desde 1993”. Esto se traduce en que los telespectadores prefieren cada vez más las plataformas *streaming* por encima de la televisión convencional. Esta diferencia de consumo es mucho más significativa entre la población joven.

Para organizar el consumo audiovisual en estas plataformas de *streaming*, a través de estos televisores inteligentes, se propone GoWatch, una aplicación interactiva basada en HTML5 para televisores basados en Samsung Smart TV. Esta aplicación permitirá a los usuarios de televisores Samsung *Smart TV* y de plataformas *streaming* establecer un control centralizado de todas las series y películas que están viendo en el momento, que querrán ver en un futuro y, además, les permitirá mantener un historial de todas las que ya han visto. También podrán guardar las series, películas o integrantes (actores, directores, guionistas, etc.) que el usuario considere que son sus favoritos para poder recurrir a estos de una manera más directa.

Para poder realizar este seguimiento personalizado de cada usuario, será imprescindible que este se registre en la aplicación. Sin embargo, habrá una serie de funcionalidades abiertas a todo usuario que disponga de la aplicación y de un televisor Samsung *Smart TV*, sin la necesidad de crearse una cuenta. De entre estas, se pueden destacar la posibilidad de visualizar las tendencias diarias de series, películas e integrantes o la búsqueda de cualquiera de estos contenidos que se acaban de enumerar y la posterior visualización de sus detalles. Será posible la obtención de todos estos datos actualizados gracias a servicios de Internet que proporcionan esta información de manera gratuita.

2 Estado del arte

Otras personas ya se habían percatado del problema plasmado en el capítulo anterior, llevar un control de todas las series y películas de tus plataformas *streaming* no es una tarea nada sencilla. Por este motivo, actualmente se pueden encontrar una inmensa variedad de herramientas que facilitan esta tarea a los usuarios.

Una característica que tienen en común todas estas soluciones ya existentes es que están disponibles solo en aplicaciones web o aplicaciones móviles. Por este motivo, la principal diferencia con la aplicación que se pretende desarrollar es que esta estará diseñada específicamente para las Samsung *Smart TV*.

TV Time [3] es una plataforma, propiedad de Whip Media, de seguimiento de series y películas que cuenta con más de 22 millones de usuarios. Obtiene los datos de su comunidad global Content Value Management (CVM) Advanced Insights [4]. Está disponible como aplicación web y como aplicación móvil para iOS y Android.

Una de sus características más destacables es que además de ser una plataforma de seguimiento, es una plataforma social, ya que sus usuarios pueden interactuar entre ellos e incluso establecer relaciones de amistad. Siguiendo su ejemplo, GoWatch también tendrá un componente social y será posible tener amigos en la aplicación y compartir recomendaciones de series y películas entre usuarios que sean amigos.

Tienen una restricción importante de que solo los usuarios registrados pueden usar su aplicación. Sin embargo, tal y como se ha mencionado en el capítulo anterior, GoWatch tendrá una serie de funcionalidades generales abiertas a cualquier usuario y para poder realizar las funciones de seguimiento personalizado ya será necesario que el usuario disponga de una cuenta.

Next Episode [5] es una aplicación móvil disponible para iOS y Android, que originalmente se creó como una página web para ayudar a organizar y a mantener un seguimiento de las películas y series favoritas de los usuarios. Obtiene los datos a través de las APIs de TheTVDB [6] y TVmaze [7] para las series y de The Movie Database [8] para las películas.

A diferencia de TV Time, Next Episode ofrece algunas funcionalidades abiertas a cualquier usuario sin necesidad de registrarse. Una de estas es la posibilidad de visualizar un calendario con los nuevos capítulos de cada día. Sin embargo, no ofrecen la posibilidad de mostrar un calendario personalizado con solo las series que sigue cada usuario. Tomando como referencia el mostrar los estrenos de capítulos, en GoWatch se dispondrá de notificaciones vía correo electrónico para avisar de la disponibilidad de nuevos capítulos, pero exclusivamente de las series que el usuario esté siguiendo.

JustWatch [9] es un motor de búsqueda para contenido en *streaming*. Disponen de una aplicación web y una aplicación móvil, disponible para iOS y Android, que se caracteriza por permitir a sus usuarios saber en qué plataformas están disponibles las series y películas según el país del usuario. Como las anteriores aplicaciones que se han nombrado en esta sección, también permite a los usuarios mantener un registro de sus series y películas. La aplicación se puede usar en su totalidad tanto sin estar registrado como si se dispone de una cuenta, pero si el usuario desea que la aplicación se sincronice en sus diferentes dispositivos debe registrarse. El usuario puede indicar las plataformas *streaming* que tiene contratadas y así tendrá la opción de filtrar el contenido por sus plataformas. Precisamente siguiendo este ejemplo, en GoWatch se podrán filtrar las series y películas que un usuario esté viendo por las plataformas que el usuario tenga registradas.

Otro punto a destacar de esta aplicación es que recolectan datos de todos sus usuarios para ofrecer recomendaciones personalizadas basadas en usuarios con gustos similares. Esto quedará totalmente fuera del alcance de GoWatch, ya que no se desarrollará ningún tipo de inteligencia artificial capaz de proporcionar esta información.

3 Implementación

La aplicación estará dirigida a las personas que en sus hogares disponen de un televisor Samsung *Smart TV* y que hacen uso frecuente de las plataformas *streaming*.

4

Los casos de uso obtenidos del diseño de los requisitos se muestran en **Fig. 1**, agrupados según el tipo de usuario que los lleva a cabo.

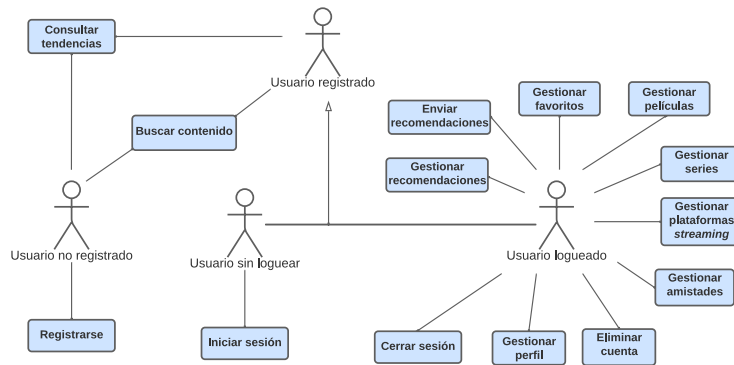


Fig. 1. Diagrama de casos de uso para GoWatch.

Para poder dar soporte a los requisitos definidos en estos casos de uso, es necesario implementar una base de datos, que solamente contiene aquella información necesaria para almacenar los usuarios y para establecer sus relaciones con los contenidos multimedia; los datos de estos contenidos (como su resumen, género, duración, etc.) se obtienen directamente de The Movie Database (TMDB). Los datos que sí que se importan en la base de datos desde TMDB se actualizan automáticamente de manera periódica para, por ejemplo, saber cuándo habrá capítulos nuevos de una serie. Solo se introducen estos datos cuando algún usuario establece una relación con ellos, para así evitar una carga inicial de todos los contenidos de TMDB.

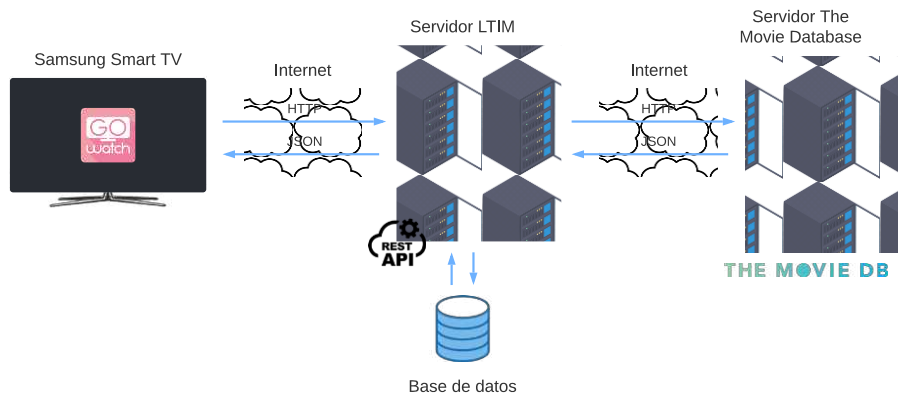


Fig. 2. Diseño arquitectónico de la aplicación.

Siguiendo la **Fig. 2**, la aplicación se comunica con el servidor a través de una API. Tanto el servidor como la API están implementados en Node.js, que obtiene y esta-

blece los datos en la base de datos de la aplicación y hace de intermediario en las peticiones a TMDb. El servidor también envía por correo electrónico las notificaciones necesarias, como el aviso de nuevos capítulos, la gestión del cambio de contraseña o los aspectos sociales y recomendaciones de la aplicación.

4 Resultados

Se ha obtenido una aplicación para Samsung Smart TV que cumple con los requisitos definidos (Fig. 1).

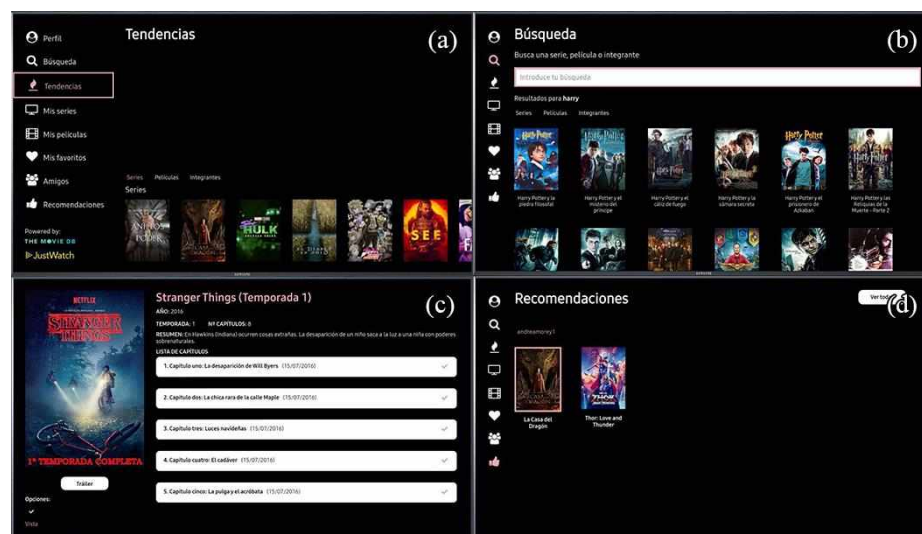


Fig. 3. Secciones tendencias (a), búsqueda (b), de seguimiento de una serie (c) y de recomendaciones (d).

A modo de ejemplo, se pueden consultar las tendencias actuales en series, películas e integrantes (Fig. 3 (a)), buscar series, películas o integrantes (Fig. 3 (b)), marcar el seguimiento del contenido audiovisual (Fig. 3 (c)), recibir y enviar recomendaciones (Fig. 3 (d)).

5 Conclusión

Siguiendo el objetivo de facilitar el seguimiento audiovisual en esta época de servicios de *streaming*, se ha obtenido esta aplicación para Samsung Smart TV que permite consultar un catálogo interactivo de contenido audiovisual (tanto series como películas), consultar los contenidos disponibles en las plataformas que se tienen contratadas, marcar el seguimiento de estos contenidos (y ser notificado al haber nuevo contenido disponible), enviar y recibir recomendaciones a través de un módulo social y marcar contenidos y personas como favoritos.

Como futuras mejoras, esta aplicación está centrada en el catálogo de las plataformas en España, así que tener disponible el catálogo de múltiples países haría que esta aplicación pudiera ser más usada. También sería interesante añadir eventos en el calendario de los usuarios para notificarlos de cuándo se estrenan nuevos contenidos que siguen. Un aspecto que tienen todas las plataformas de *streaming* es un módulo de recomendaciones a partir de los contenidos vistos; ya que disponemos de estos datos, podríamos recomendar contenido audiovisual basado en los contenidos visualizados o en los contenidos y personas marcados como favoritos. Finalmente, al disponer de un módulo social, se podría ampliar e incorporar una funcionalidad similar a la que ha introducido Spotify recientemente, “Actividad de amigos” [10], que mostraría qué contenidos están consumiendo los amigos en ese momento.

Referencias

- [1] AIMC, “Marco General de los Medios en España,” 2022.
- [2] C. Martínez, “Las plataformas causan la mayor caída del consumo televisivo en 30 años,” 2022.
- [3] “TV Time | Your TV shows calendar (TVShow Time).” <https://www.tvtime.com/> (accessed Nov. 11, 2022).
- [4] “Whip Media | Reimagined Content Licensing for Today.” <https://www.whipmedia.com/#> (accessed Nov. 11, 2022).
- [5] “Track the TV shows and movies you watch - Next Episode.” <https://next-episode.net/> (accessed Nov. 11, 2022).
- [6] “API - TheTVDB.com.” <https://thetvdb.com/api-information> (accessed Nov. 11, 2022).
- [7] “TV API | TVmaze - Add TV information to your website or app.” <https://www.tvmaze.com/api> (accessed Nov. 11, 2022).
- [8] “The Movie Database (TMDB).” <https://www.themoviedb.org/> (accessed Nov. 11, 2022).
- [9] “JustWatch - The Streaming Guide.” <https://www.justwatch.com/> (accessed Nov. 11, 2022).
- [10] “Friend Activity - Spotify.” <https://support.spotify.com/us/article/friend-activity/> (accessed Nov. 11, 2022).

Revisão sistemática da literatura sobre o mapeamento de emoções a partir de leituras neurais para fruição de conteúdos audiovisuais

Daniel de Queiroz Cavalcanti¹[0000-0002-7875-7076], Felipe Melo¹[0000-0003-3930-5760],
Thiago Silva¹[0000-0003-1689-1736] and Valdecir Becker¹[0000-0002-0639-3875]

¹ Laboratory of Interaction and Media, Informatics Center,
Federal University of Paraíba, João Pessoa - PB, Brazil
contato@lim.ci.ufpb.br

Abstract. Este artigo apresenta uma revisão sistemática da literatura (RSL) sobre conceitos, mapeamento e análise de emoções dos indivíduos durante a fruição de conteúdos audiovisuais, utilizando interfaces cérebro máquina. Baseado no protocolo PICO, a coleta dos dados foi feita nas bases ACM, Springer e IEEE, considerando artigos publicados entre 2013 e 2021. Foram identificados 151 artigos. Destes, 95 foram lidos, e a partir dos critérios de inclusão e exclusão, 56 foram analisados em profundidade. Os dados foram analisados e organizados conforme as tecnologias e processo de pesquisa. Como principais conclusões, descobriu-se que, para analisar emoções existem níveis com focos variados, especialmente as avaliações macro e micro. No primeiro caso, o resultado é determinístico, como uma emoção padrão; no segundo, a análise é feita a partir de uma avaliação específica ou uma combinação das avaliações, que variam entre diferentes indivíduos. Além disso, foram identificadas pesquisas para indução do humor e ativação da linguagem com base em ressonância magnética. Pesquisadores também propuseram métodos para medir as diferenças entre os dois hemisférios. Apesar dos avanços recentes, esta RSL mostra que o tema é um tópico de interesse científico, especialmente no que se refere à utilização da assimetria do cérebro para melhorar o desempenho do reconhecimento de padrões emocionais, principalmente a partir da eletroencefalografia.

Keywords: RSL, EEG, Fruição Audiovisual.

1 Introdução

Interfaces cérebro-máquina (ICM) se tornaram comuns na interação humano-computador. Novas formas de interação foram incorporadas, inovando na troca de informações entre indivíduos e sistemas computacionais. Uma ICM que tem se tornado cada vez mais comum é baseada na eletroencefalografia (EEG), ou seja, na leitura das ondas neurais e na consequente geração de dados binários para uso por computadores.

Esse tipo de interface é mais comum, e disseminada, na área da saúde. Pesquisas têm apontado grande potencial no tratamento de traumas, tanto físicos quanto psíquicos. No entanto, poucos estudos foram identificados sobre o uso de EEG como ICM em outras

áreas do conhecimento, como entretenimento e fruição de conteúdos audiovisuais. Apesar de headsets estarem disponíveis comercialmente com uma grande variedade de formatos e preços, há uma limitação de estudos sobre o uso dessa tecnologia para mapeamento de emoções, gostos e relações subjetivas com conteúdos audiovisuais.

Dentro desse contexto, foi realizada uma revisão sistemática da literatura (RSL) para identificar pesquisas sobre o tema e responder a duas perguntas centrais: Como o tema emoção é abordado nas pesquisas utilizando ICMs? Há pesquisas que apontam para a utilização de ICMs baseadas em EEG para identificação de emoções durante o processo de fruição audiovisual? Visando uma melhor compreensão do tema, e delimitar de forma mais objetiva o que é pesquisa envolvendo emoção e o que não é, foi necessário incluir na revisão artigos relacionados a conceitos atuais sobre emoção. Dessa forma, o escopo da primeira pergunta foi expandido.

Analisando as três bases de dados mais importantes para a área da interação humano-computador, ACM, Springer e IEEE, e aplicando os critérios de inclusão e exclusão, foram identificados 56 artigos sobre o tema. Uma síntese destes trabalhos é apresentada neste artigo. Como conclusão principal, percebeu-se que a identificação das emoções nos seres humanos é feita principalmente com ressonância magnética, técnica esta que permite análises bem complexas. A maior parte das pesquisas aborda as avaliações nos níveis macro e micro. No nível macro o resultado é determinístico, como uma emoção padrão; já no nível micro, a análise é feita a partir de uma avaliação específica ou uma combinação das avaliações, que variam entre diferentes indivíduos.

2 Métodos

As RSLs são consideradas como um método de pesquisa chave para apoiar a pesquisa baseada em evidências [60]. Esta revisão inclui artigos de três bases de dados diferentes e classifica os resultados dos problemas e soluções propostas pelos artigos analisados. Para esta RSL, adotamos o protocolo PICO (Population, Intervention, Comparison and Outcomes) para organização inicial. A população compreende estudos primários sobre teorias e métodos de usos de EEG como ICM, além dos conceitos e análise de emoções. A intervenção indica conceitos, teorias ou relatórios sobre uso, desenvolvimento e avaliação de ICMs. Para Comparação e Resultados buscou-se verificar quais estudos estavam diretamente envolvidos com o uso ou desenvolvimento de EEG como ICM e análise de emoções ou tinham potencial contribuição para o desenvolvimento de ferramentas similares. A partir desses tópicos, foram levantadas as seguintes questões: o que se discute sobre o desenvolvimento e uso de EEG como ICM, quais são as principais tecnologias, métodos e conceitos utilizados para o mapeamento de emoções e identificação de afinidades com os conteúdos.

A *string* de busca foi adaptada em função dos diferentes mecanismos de busca das bibliotecas, contemplando as seguintes palavras-chave: affective computing, neuroscience, man-machine interaction, emotions, shape emotions, Rendering Emotions, Emotion Tracking, emotion recognition, EEG, emotion detection, electroencephalography, emotion detection using electroencephalography; affective computing neuromarketing, artificial intelligence affective computing.

Os critérios de inclusão que foram utilizados no mapeamento: (i) Conter as palavras-chave da string de busca no título ou resumo; (ii) Possuir versão disponível online; (iii) Em caso de artigos que tratavam da mesma pesquisa, apenas o mais recente foi incluído.

A partir dos artigos encontrados, definiu-se os critérios de exclusão para manter assim o foco da pesquisa: (i) Artigos cujo foco principal não era o conceito, a análise de emoções ou o uso de EEG como ICM; (ii) Artigos específicos sobre a área da saúde; (iii) Mapeamentos e revisões sistemáticas que apenas usavam EEG e ICM como exemplos. Aplicando a string de busca nas três bases de dados, foram localizados 151 artigos. Seguindo os critérios de inclusão e exclusão em cada artigo, restaram 56 artigos, que compõem o presente artigo. Destes, 37 são da ACM Digital Library, 13 da IEEE e 6 da Springer (Fig. 1).

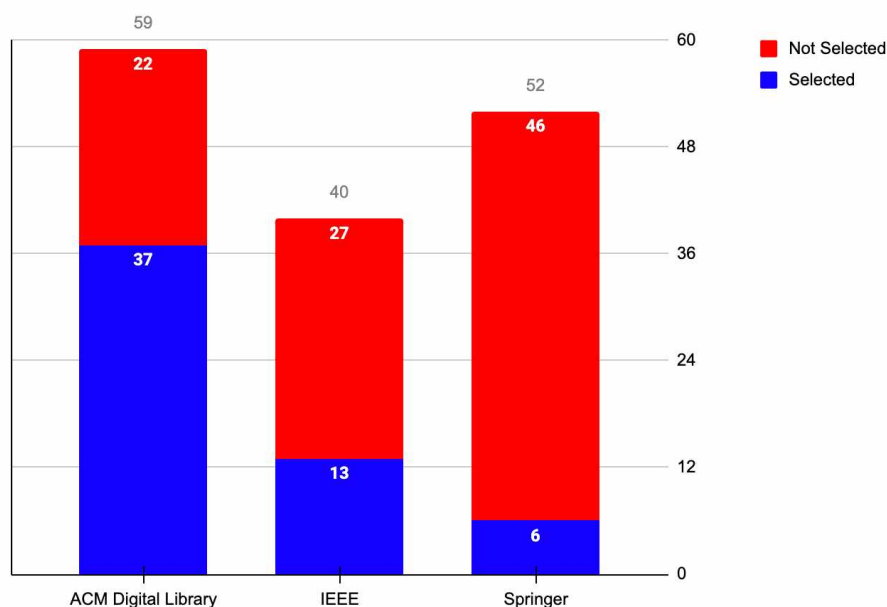


Fig. 1. Proporção de aproveitamento dos artigos [Autoria própria].

Para a análise das informações e organização dos dados, os artigos foram agrupados em categorias: conceitos e teorias sobre emoções, técnicas de análise, tecnologias de mapeamento. O desmembramento dessas categorias é descrito a seguir.

3 Resultados, análise e discussão

Do ponto de vista da neurociência, existem regiões do córtex cerebral que estão intimamente relacionadas com as emoções [1]. Tal constatação nos fornece uma maneira potencial de decodificar a emoção gravando os sinais do cérebro humano dessas regiões. Consequentemente, em tese, também é possível identificar significados da emoção. Uma forma de fazer isso é acoplado os eletrodos de EEG no couro cabeludo para

registro das atividades neurais do cérebro. Outra forma comum é através de imagens por ressonância magnética funcional, onde se detecta variações no fluxo sanguíneo em resposta à atividade neural. Os dados gerados podem ser usados para reconhecer as emoções do ser humano.

Dentro dessa perspectiva, cientistas têm realizado pesquisas de indução do humor e ativação da linguagem com base em ressonância magnética (fMRI), descobrindo que a tristeza e felicidade são processados por cada hemisfério [2,3,5,7,8,24,25,31,37,42,47]. Além disso, pesquisadores [2,5,33,34,35,42] também propuseram métodos para medir as diferenças entre os dois hemisférios, que foram utilizados para a detecção de depressão e monitoramento de EEG. No entanto, ainda é um tópico interessante de como utilizar a propriedade de assimetria do cérebro para melhorar o desempenho de reconhecimento da emoção.

Conceitualmente, a análise das emoções pode ser dividida em dois níveis, macro e micro [2]. No nível macro o resultado é determinístico, como uma emoção padrão. Além disso, pesquisadores [2,24] defendem o emergentismo, assumindo que a combinação da avaliação de elementos no processo recursivo está se desdobrando ao longo do tempo. Essa reação resultante forma emoções emergentes que são mais do que a soma de seus constituintes e mais do que instantâneo de categorias rígidas.

Em contrapartida, as microemoções foram descobertas pelos pesquisadores Ekman e Friesen em 1969, ao identificarem as emoções ocultas de indivíduos [57]. As microemoções estão diretamente relacionadas a microexpressões e compreendem um sentimento involuntário, ou não consciente, que vem à tona quando as pessoas tentam esconder ou reprimir suas emoções.

A detecção de microexpressões ajuda a entender o real estado emocional das pessoas, mesmo que essa detecção esteja descorrelacionada com a percepção consciente [57]. Essa análise de estado verdadeiro ou de informações ocultas têm um grande impacto sobre os campos de diagnóstico psicológico, fisiológico, e ambientes interrogativos. Além disso, também tem amplas aplicações em os cuidados de saúde para o tratamento do autismo e determinando clientes estado médico, durante os cuidados de saúde. Isso motivou a estudo e análise de emoções ocultas em sinais de fala, por exemplo.

Visando relacionar as emoções com expressões físicas, os autores [58] fizeram pesquisas na detecção de micro e macroexpressão facial em longas sequências de vídeo usando quadro de referência orientado temporal. Para isso, eles utilizaram o método leave-one-out, ou seja, foi adotado um método que realiza N cálculos de erro, um para cada dado. Apesar de apresentar uma investigação completa sobre a variação do modelo em relação aos dados utilizados, esse método possui um alto custo computacional, sendo indicado apenas para situações em que poucos dados estão disponíveis. Sabendo disso, foi proposto uma análise das micro e macroexpressões, através de vídeos. A solução encontrada é baseada no aprendizado profundo que, em vez de rastrear o movimento diferencial, comparou cada quadro do vídeo usando um modelo convolucional. Dessa forma, com dois quadros de referência locais temporalmente, foram mostrados de acordo com a duração das micro e macroexpressões calculadas do usuário. Entretanto, houve melhor desempenho em um conjunto de dados de vídeos longos SAMM de alta taxa de quadros (200 fps) (SAMM-LV) do que em uma taxa de quadros baixa

(30 fps) (CAS(ME)2) e foi possível alcançar o F1-Score de 0,1531 como resultado de linha de base.

Outro conceito relevante identificado nesta RSL é o modelo de processo de componente (CPM) [4]. Trata-se de uma visão para explicar emoções como resultados emergentes de dimensões subjacentes, adicionando uma camada de mecanismo na base da experiência emocional. Este modelo assume uma série de avaliações cognitivas em diferentes níveis de processamento que introduz mais flexibilidade e pode explicar diferenças individuais e contextuais em respostas aos mesmos estímulos. Além disso, tal modelo pode acomodar a emoção não prototípica e mista/difusa. De acordo com o modelo de processo componente da emoção, as mudanças de subsistemas em um organismo são impulsionadas por avaliações que refletem a avaliação subjetiva de um evento e ativa alterações em outros subsistemas. Neste modelo, cada experiência emocional surge de mudanças coordenadas em cinco componentes: 1) o componente de avaliação, que implica avaliar o evento/situação em relação a significado, implicações para os objetivos, potencial de enfrentamento, novidade, e compatibilidade com as normas; as avaliações são tiradas de memória e representações de objetivos e examinar a relevância do evento e suas consequências para o bem-estar do organismo. As mudanças neste componente acionam mudanças nos outros quatro componentes principais; 2) uma motivação componente que define mudanças nas tendências de ação (por exemplo, lutar, fugir ou congelar) e se prepara para a(s) resposta(s) apropriada(s); 3) um componente fisiológico que engloba mudanças na atividade autonômica periférica (por exemplo, alterações na frequência cardíaca ou frequência respiratória) e ocorre com base nos resultados da avaliação e mudanças motivacionais associadas; 4) um componente de expressão que envolve mudanças no comportamento motor expressivo, como expressão facial, gestos/posturas corporais e também é ativado semelhante ao componente de fisiologia como resultado da avaliação e componentes de motivação; e finalmente 5) um componente de sentimento que reflete a experiência consciente associada às mudanças em todos os outros componentes, geralmente descritos por pessoas com alguns rótulos categóricos (como raiva, felicidade, tristeza e assim por diante) .

Essa abordagem vem sendo aplicada ao desenvolvimento de agentes artificiais, cada vez mais presentes em nosso meio social, com um número crescente desses agentes aparecendo em contextos de hospitalidade, cuidado e educação [5]. Um exemplo é o desenvolvimento de robôs humanoides. As pesquisas têm apontado que, para maximizar a qualidade das interações sociais entre humanos e agentes artificiais, é importante que o agente artificial não responda apenas às emoções expressas pelo agente humano, mas também seja capaz de expressar emoções em si. À medida que o interesse e o investimento em robótica social continuam para crescer, desenvolver agentes artificiais com esse tipo de capacidade emocional é um requisito fundamental para a robótica verdadeiramente social agentes. Enquanto o componente emocional dos agentes artificiais tem sido negligenciado até recentemente, quando investigações empíricas começaram a mudar o foco. Pesquisas emergentes documentam como os robôs são expressivos, podendo ser classificados como mais simpáticos e humanizados, levando a um maior engajamento e interações mais prazerosas. A aceitação e cooperação com um robô depende da correspondência entre a situação e o comportamento emocional dele [5].

Durante as últimas décadas, o reconhecimento de emoções tem desempenhado um papel importante na interação humana, despertando cada vez mais interesse e atenção [6]. No entanto, boa parte dos trabalhos nesta área tem focado na análise das emoções representadas ou estereotipadas. Sabendo disso, na maioria dos estudos [2,6,9] sobre emoções ocorre uma classificação em algumas emoções básicas discretas [9] (por exemplo, felicidade, tristeza, surpresa, medo, raiva e nojo). Embora muitos resultados de reconhecimento promissores foram alcançados recentemente, o estado da arte atual ainda não atende todas as necessidades da vida real, porque existe estados emocionais não básicos, sutis e bastante complexos, que não podem ser totalmente expressos, ou classificados, por um rótulo ou uma categoria de emoção [10].

Focado nisso, pesquisadores [2,9,10,11,12,20,21,39,40,44,46] argumentam que a dimensão da abordagem de modelagem de emoções é mais adequada para expressar as emoções complexas. Dessa forma, ocorre uma tentativa de aprendizagem da emoção em um espaço de multidimensional, em vez de alguns conceitos básicos discretos. Por exemplo, a dimensão de excitação refere-se à emoção como excitada ou apática. A dimensão de valência refere-se para quão positiva ou negativa é a emoção. A dimensão de dominância refere-se ao grau de poder ou senso de controle sobre a emoção [13,14,15]. Assim, os vários estados emocionais localizados em diferentes posições e suas semelhanças e diferenças podem ser expressas por suas distâncias neste espaço. Diante disso, outras pesquisas [9,16,17,18] sobre este tema aumentaram o escopo, não se restringindo apenas a detectar várias emoções prototípicas, mas também reconhecer os estados da emoção a cada momento. Dessa forma, é possível trabalhar numa direção sutil, gerando interpretações contínuas e específicas do contexto de monitores gravados em configurações do mundo real.

Outro elemento relevante identificado pela presente RSL está no reconhecimento da emoção na dimensão auditiva, visual e fisiológica. Sediada na arquitetura de previsão Long Short-Term Memory Recurrent Networks (LSTM-RNN) [9], que pode ser explicada como o estado classificador de arte para reconhecimento dimensional de emoções [58,59], usam-se duas técnicas introduzidas no reconhecimento dimensional de emoções: 1. Perda insensível a ϵ : a perda insensível a ϵ é utilizada como a perda função para treinar a rede neural. O ruído do rótulo é um problema inevitável para a emoção dimensional. Comparado com função de perda ao quadrado, que é a perda mais amplamente utilizada função para reconhecimento de emoção de dimensão, insensível a ϵ a perda é mais robusta para os ruídos da etiqueta. Além disso, a perda ϵ insensível pode ignorar pequenos erros para ficar mais forte correlação entre as previsões e rótulos.

A LSTM-RNN é uma das técnicas de aprendizado de máquina de última geração em reconhecimento dimensional de emoções. Tem a capacidade de incorporar conhecimento sobre como as emoções geralmente evoluem ao longo do tempo, de modo que as estimativas de emoção inferidas são produzidas sob consideração de uma quantidade ideal de contexto [16]. Dessa forma, o desafio gerado pelo problema da regressão para emoção dimensional é simplificado para um problema de classificação positiva e negativa. Combinando as modalidades de áudio e visual, [17] obtém o estado da arte de atuação.

Para o problema de regressão, focado diretamente no reconhecimento de emoção dimensional, Jia-Hao Hsu et al. Várias pesquisas têm investigado a LSTM-RNN para

modalidades de áudio, visuais e fisiológicas baseadas reconhecimento dimensional de emoções. Os dados assíncronos entre classificações contínuas e dados sobre a análise do tamanho da janela para dimensões de emoção e fusão multimodal compõe a centro da análise em [18]. Já Damien Dupré et al. [19] também utilizam LSTM-RNN para analisar o problema de reconhecimento dimensional de emoções. A obtenção de alta concordância interobservadora é um dos principais desafios na anotação de dados de emoção, especialmente para análise da emoção dimensional [22,23]. Assim, o ruído do rótulo é um problema inevitável atualmente. Michel et al. [26] buscam minimizar o ruído do rótulo para calcular as classificações médias de todos os avaliadores [27,28,29,55,56,59], centralizando os dados de diferentes avaliadores de acordo com o valor médio de todos as entradas. O resultado foi uma combinação dessas classificações ponderadas linearmente por seus respectivos acordos entre avaliadores. A concordância é medida pelo coeficiente de correlação de concordância (CCC).

Além disso, [30,32,36] adicionaram outros recursos ao rótulo em determinadas janelas de tempo. Por calcular a média dos rótulos na janela, o rótulo é suavizado e o ruído diminui até certo ponto. Já os pesquisadores [2,6,9,27,28,29,55,56,59] buscaram minimizar os ruídos de anotações de diferentes maneiras. No entanto, quando os rótulos com ruído são dados, raramente considera-se este problema a partir da perda de função. Além disso, [38,40,41] agregaram o agrupamento temporal, descrita como uma técnica de reconhecimento de emoção dimensional utilizando a função na rede de crenças profundas. Essa técnica permite a modelagem temporal na rede direta, gerando resultados bem competitivos [43,44,45]. Nesse contexto, é utilizada a função de agrupamento temporal antes da camada do LSTM [46,48,49], a modelagem temporal em escala curta [50,51,52], visando aumentar a diversidade dos recursos alimentados em camadas de previsão para frente [53,54,55,56].

4 Conclusão

Este artigo descreveu os resultados de uma revisão sistemática da literatura (RSL) sobre os conceitos pertinentes, técnicas de mapeamento e de análise de emoções dos indivíduos a partir de interfaces cérebro máquina. A pesquisa foi baseada no protocolo PICO, com coleta dos dados nas bases ACM, Springer e IEEE, resultando em 56 artigos. A pesquisa objetiva dois pontos centrais. O primeiro buscava identificar como a emoção era abordada, de forma ampla, nas pesquisas utilizando ICMs na fruição de conteúdos audiovisuais. Foram identificadas diferentes abordagens, consideradas relevantes dentro do contexto da pesquisa. A classificação das emoções, suas origens e formas de mapeamento baseadas em fMRI se fazem presentes na maioria dos 56 artigos estudados a fundo. Além disso, técnicas baseadas em inteligência artificial e aprendizado de máquina também foram identificadas. Tais pesquisas apontam para uma abordagem empírica visando desenvolver e emular emoções em sistemas computacionais.

Além disso, um segundo objetivo da RSL visava identificar pesquisas que utilizassem a eletroencefalografia como interface cérebro-máquina para mapear emoções durante o processo de fruição audiovisual. Considerando a metodologia adotada, as três bases científicas pesquisadas e o espaço temporal, não foram identificadas pesquisas

específicas sobre EEG como ICM para identificar, ou mapear, emoções durante a fruição de conteúdos audiovisuais.

Como limitações e melhorias para os trabalhos futuros, é preciso ressaltar a ausência de resultados nas áreas da computação afetiva e neurociência afetiva. São duas áreas de potenciais contribuições para o tema, mas com resultados inexpressivos nos métodos utilizados pela presente RSL. É possível que estudos pertinentes não tenham sido mapeados devido a idiosincrasias dessas áreas, especialmente no que se refere a conceitos e palavras-chave utilizadas. Outro fator limitante está no fato de determinados artigos terem apenas os resumos publicados online, gerando desclassificação pelos critérios de exclusão. Finalmente, a quantidade de pesquisadores envolvidos na revisão sistemática (três) também foi um fator limitante, pois mais participantes poderiam auxiliar na tomada de decisão, diminuído alguma subjetividade ainda presente na aplicação dos critérios de inclusão e exclusão.

Para trabalhos futuros, sugere-se um mapeamento sistemático (Scoping Review), um tipo de pesquisa exploratória que mapeia sistematicamente a literatura sobre um tópico, identificando conceitos-chave, teorias e fontes de evidências que informam a prática na área. Trata-se de uma abordagem mais ampla do que a RSL, permitindo inclusão de pesquisas fora das bases centrais, mas mantendo o rigor científico e a validade da interpretação dos dados.

Referências

1. LI, Yang et al. A bi-hemisphere domain adversarial neural network model for EEG emotion recognition. *IEEE Transactions on Affective Computing*, v. 12, n. 2, p. 494-504, 2018.
2. Dattada, Vamsi Vijay Mohan; Jeevan, M. Analysis of concealed anger emotion in a neutral speech signal. In: 2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER). IEEE, 2019. p. 1-5.
3. Alimuradov, Alan K.; Tychkov, Alexander Yu; CHURAKOV, Pyotr P. A novel approach to speech signal segmentation based on empirical mode decomposition to assess human psycho-emotional state. In: 2019 3rd School on Dynamics of Complex Networks and their Application in Intellectual Robotics (DCNAIR). IEEE, 2019. p. 9-12.
4. Mohammadi, Gelareh; Vuilleumier, Patrik. A multi-componential approach to emotion recognition and the effect of personality. *IEEE Transactions on Affective Computing*, 2020.
5. Hortensius, Ruud; Hekele, Felix; Cross, Emily S. The perception of emotion in artificial agents. *IEEE Transactions on Cognitive and Developmental Systems*, v. 10, n. 4, p. 852-864, 2018.
6. Fatemeh Shahrabi Farahani; Mansour Sheikhan; Ali Farrokhi. A fuzzy approach for facial emotion recognition. 2013. 13th Iranian Conference on Fuzzy Systems (IFSC). doi: 10.1109/IFSC.2013.6675597
7. Raj Rakshit, V Ramu Reddy e Parijat Deshpande. 2016. Emotion detection and recognition using HRV features derived from photoplethysmogram signals. Em Anais do 2º workshop sobre Representações e Modelagem de Emoções para Sistemas Companheiros (ERM4CT '16). Association for Computing Machinery, Nova York, NY, EUA, Artigo 2, 1-6. <https://doi.org/10.1145/3009960.3009962>
8. Hristijan Gjoreski, Ifigeneia I. Mavridou, Mohsen Fatoorechi, Ivana Kiprijanovska, Martin Gjoreski, Graeme Cox, and Charles Nduka. 2021. EmteqPRO: Face-mounted Mask for

- Emotion Recognition and Affective Computing. Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers. Association for Computing Machinery, New York, NY, USA, 23–25. <https://doi.org/10.1145/3460418.3479276>
9. Linlin Chao, Jianhua Tao, Minghao Yang, Ya Li, and Zhengqi Wen. 2015. Long Short Term Memory Recurrent Neural Network based Multimodal Dimensional Emotion Recognition. In Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge (AVEC '15). Association for Computing Machinery, New York, NY, USA, 65–72. <https://doi.org/10.1145/2808196.2811634>
 10. Subhasmita Sahoo; Aurobinda Routray. 2019. MFCC feature with optimized frequency range: An essential step for emotion recognition. 2016 International Conference on Systems in Medicine and Biology (ICSMB). <https://doi.org/10.1109/ICSMB.2016.7915112>
 11. S. A. Hassan, S. Akbar, A. Rehman, T. Saba, H. Kolivand and S. A. Bahaj, "Recent Developments in Detection of Central Serous Retinopathy Through Imaging and Artificial Intelligence Techniques—A Review," in IEEE Access, vol. 9, pp. 168731-168748, 2021, doi: 10.1109/ACCESS.2021.3108395.
 12. G. R. Djavanshir, X. Chen and W. Yang, "A Review of Artificial Intelligence's Neural Networks (Deep Learning) Applications in Medical Diagnosis and Prediction," in IT Professional, vol. 23, no. 3, pp. 58-62, 1 May-June 2021, doi: 10.1109/MITP.2021.3073665.
 13. Yoichi Yamazaki; Michiya Yamamoto; Noriko Nagata, "Estimation of Emotional State in Personal Fabrication: Analysis of Emotional Motion Based on Laban Movement Analysis". 2017 IEEE International Conference on Culture and Computing (Culture and Computing), doi: 10.1109/Culture.and.Computing.2017.33
 14. Yelin Kim; Emily Mower Provost, "ISLA: Temporal Segmentation and Labeling for Audio-Visual Emotion Recognition," IEEE Transactions on Affective Computing (Volume: 10, Issue: 2, 01 April-June 2019). doi: 10.1109/TAFFC.2017.2702653
 15. Lanxin Sun; JunBo Dai; Xunbing Shen, "Facial emotion recognition based on LDA and Facial Landmark Detection,"2021 2nd International Conference on Artificial Intelligence and Education (ICAIE), doi: 10.1109/ICAIE53562.2021.00020.
 16. G. Valenza, L. Citi, A. Lanatà, E. P. Scilingo and R. Barbieri, "A nonlinear heartbeat dynamics model approach for personalized emotion recognition," 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 2579-2582, doi: 10.1109/EMBC.2013.6610067.
 17. Dong Hwa Kim and Dae Sung Seo. 2019. Vector based 3D Emotion Expression for Emotion Robot. In Proceedings of the 5th International Conference on Mechatronics and Robotics Engineering (ICMRE'19). Association for Computing Machinery, New York, NY, USA, 113–117. <https://doi.org/10.1145/3314493.3314499>
 18. Jia-Hao Hsu; Chung-Hsien Wu, "Attentively-Coupled Long Short-Term Memory for Audio-Visual Emotion Recognition" in 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC).
 19. Damien Dupré; Anna Tcherkassof; Michel Dubois, "Emotions triggered by innovative products: A multi-componential approach of emotions for user experience tools," in IEEE Transactions on Affective Computing, vol. 12, no. 5, pp. 1120-1129, 21-24 Sept. 2015, doi: 10.1109/ACII.2015.7344657
 20. M. Danai Styliani. 2016. An Affective Service based on Multi-Modal Emotion Recognition, using EEG enabled Emotion Tracking and Speech Emotion Recognition.2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia). doi: 10.1109/ICCE-Asia49877.2020.9277291

21. Smitha Engoor; S. SendhilKumar; C. Hepsibah Sharon; G.S. Mahalakshmi. Occlusion-aware Dynamic Human Emotion Recognition Using Landmark Detection. 2020. 6th International Conference on Advanced Computing and Communication Systems (ICACCS). doi: <https://doi.org/10.1109/ICACCS48705.2020.9074318>
22. Sneha Lukose; Savitha S. Upadhy. 2017. Music player based on emotion recognition of voice signals. International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT).doi: 10.1109/ICICT1.2017.8342835
23. De'Aira Bryant and Ayanna Howard. 2019. A Comparative Analysis of Emotion-Detecting AI Systems with Respect to Algorithm Performance and Dataset Diversity. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19). Association for Computing Machinery, New York, NY, USA, 377–382. <https://doi.org/10.1145/3306618.3314284>
24. Claudia Faita, Federico Vanni, Camilla Tanca, Emanuele Ruffaldi, Marcello Carrozzino, and Massimo Bergamasco. 2016. Investigating the process of emotion recognition in immersive and non-immersive virtual technological setups. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16). Association for Computing Machinery, New York, NY, USA, 61–64. <https://doi.org/10.1145/2993369.2993395>
25. M. L. Menezes, A. Samara, L. Galway, A. Sant'anna, A. Verikas, F. Alonso-Fernandez, H. Wang, and R. Bond. 2017. Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset. *Personal Ubiquitous Comput.* 21, 6 (December 2017), 1003–1013. <https://doi.org/10.1007/s00779-017-1072-7>
26. Linlin Chao, Jianhua Tao, Minghao Yang, Ya Li, and Zhengqi Wen. 2014. Multi-scale Temporal Modeling for Dimensional Emotion Recognition in Video. In Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge (AVEC '14). Association for Computing Machinery, New York, NY, USA, 11–18. <https://doi.org/10.1145/2661806.2661811>
27. Hao Jiang, Zhigang Deng, Mingliang Xu, Xiangjun He, Tianlu Mao, and Zhaoqi Wang. 2018. An emotion evolution based model for collective behavior simulation. In Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '18). Association for Computing Machinery, New York, NY, USA, Article 10, 1–6. <https://doi.org/10.1145/3190834.3190844>
28. Ali Harimi; Ali Shahzadi; Alireza Ahmadyard. Recognition of emotion using non-linear dynamics of speech. 2014. 7'th International Symposium on Telecommunications (IST'2014). doi: 10.1109/ISTEL.2014.7000745.
29. Jiaxin Ma, Hao Tang, Wei-Long Zheng, and Bao-Liang Lu. 2019. Emotion Recognition using Multimodal Residual LSTM Network. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 176–183. <https://doi.org/10.1145/3343031.3350871>
30. Mingmin Zhao, Fadel Adib, and Dina Katabi. 2018. Emotion recognition using wireless signals. *Commun. ACM* 61, 9 (September 2018), 91–100. <https://doi.org/10.1145/3236621>
31. Zhengwei Huang, Ming Dong, Qirong Mao, and Yongzhao Zhan. 2014. Speech Emotion Recognition Using CNN. In Proceedings of the 22nd ACM international conference on Multimedia (MM '14). Association for Computing Machinery, New York, NY, USA, 801–804. <https://doi.org/10.1145/2647868.2654984>
32. T. Liogienė and G. Tamulevičius, "SFS feature selection technique for multistage emotion recognition," 2015 IEEE 3rd Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), 2015, pp. 1-4, doi: 10.1109/AIEEE.2015.7367299.

33. G. Wei, L. Jian and S. Mo, "Multimodal (Audio, Facial and Gesture) based Emotion Recognition challenge," 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), 2020, pp. 908-911, doi: 10.1109/FG47880.2020.00142.
34. D. Sokolov and M. Patkin, "Real-Time Emotion Recognition on Mobile Devices," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, pp. 787-787, doi: 10.1109/FG.2018.00124.
35. T. Keshari and S. Palaniswamy, "Emotion Recognition Using Feature-level Fusion of Facial Expressions and Body Gestures," 2019 International Conference on Communication and Electronics Systems (ICCES), 2019, pp. 1184-1189, doi: 10.1109/ICCES45898.2019.9002175.
36. Cynthia Joesph; A. Rajeswari; B. Premalatha; C. Balapriya. Implementation of physiological signal based emotion recognition algorithm. 2020 IEEE 36th International Conference on Data Engineering (ICDE). doi: 10.1109/ICDE48307.2020.9153878
37. V. Gonuguntla and J. -H. Kim, "EEG-Based Functional Connectivity Representation using Phase Locking Value for Brain Network Based Applications," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 2853-2856, doi: 10.1109/EMBC44109.2020.9175397.
38. Elif Gümüslü, Duygun Erol Barkana, and Hatice Köse. 2021. Emotion Recognition using EEG and Physiological Data for Robot-Assisted Rehabilitation Systems. In Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion). Association for Computing Machinery, New York, NY, USA, 379–387. <https://doi.org/10.1145/3395035.3425199>
39. Chuqiao Wei; Hua Zhang; Liang Ye; Fanchao Men. 2020. A school bullying detecting algorithm based on motion recognition and speech emotion recognition. 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI). doi: <https://doi.org/10.1109/ICHCI51889.2020.00066>
40. Zhen Gao and Shangfei Wang. 2015. Emotion Recognition from EEG Signals using Hierarchical Bayesian Network with Privileged Information. In Proceedings of the 5th ACM International Conference on Multimedia Retrieval (ICMR '15). Association for Computing Machinery, New York, NY, USA, 579–582. <https://doi.org/10.1145/2671188.2749364>
41. T. Yang, W. Huang and K. K. Toe, "Statistical modeling on motion trajectories for robotic laparoscopic surgery," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2017, pp. 4347-4350, doi: 10.1109/EMBC.2017.8037818.
42. Sourabh Prajapati, C. L. Shrinivasa Naika, Shashi Shekhar Jha, and Shivashankar B. Nair. 2013. On Rendering Emotions on a Robotic Face. In Proceedings of Conference on Advances In Robotics (AIR '13). Association for Computing Machinery, New York, NY, USA, 1–7. <https://doi.org/10.1145/2506095.2506151>
43. E. Bekele et al., "Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD)," 2016 IEEE Virtual Reality (VR), 2016, pp. 121-130, doi: 10.1109/VR.2016.7504695.
44. R. Gill and J. Singh, "A Review of Neuromarketing Techniques and Emotion Analysis Classifiers for Visual-Emotion Mining," 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), 2020, pp. 103-108, doi: 10.1109/SMART50582.2020.9337074.
45. Samer Schaat, Aleksandar Miladinović, Stefan Wilker, Stefan Kollmann, Stephan Dickert, Erdem Geveze, and Verena Gruber. 2015. Emotion in Consumer Simulations for the Development and Testing of Recommendations for Marketing Strategies. In Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Systems 2015 (EMPIRE '15).

- Association for Computing Machinery, New York, NY, USA, 25–32. <https://doi.org/10.1145/2809643.2809649>
46. Subhashini Sivagnanam, Kenneth Yoshimoto, Nicholas T. Carnevale, and Amit Majumdar. 2018. The Neuroscience Gateway: Enabling Large Scale Modeling and Data Processing in Neuroscience. In Proceedings of the Practice and Experience on Advanced Research Computing (PEARC '18). Association for Computing Machinery, New York, NY, USA, Article 52, 1–7. <https://doi.org/10.1145/3219104.3219139>
 47. Jérôme Guzzi, Alessandro Giusti, Luca M. Gambardella, and Gianni A. Di Caro. 2018. A model of artificial emotions for behavior-modulation and implicit coordination in multi-robot systems. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '18). Association for Computing Machinery, New York, NY, USA, 21–28. <https://doi.org/10.1145/3205455.3205650>
 48. Suchitra; Suja P.; Shikha Tripathi. 2016. Real-time emotion recognition from facial images using Raspberry Pi II. 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN). doi: 10.1109/SPIN.2016.7566780
 49. Suchitra; Suja P.; Shikha Tripathi. Taranpreet Singh Saini, Mangesh Bedekar, and Saniya Zahoor. 2017. Circle of Emotions in Life: Emotion Mapping in 2Dimensions. In Proceedings of the 9th International Conference on Computer and Automation Engineering (ICCAE '17). Association for Computing Machinery, New York, NY, USA, 83–88. <https://doi.org/10.1145/3057039.3057046>
 50. Kata, G., Poleszak, W.: Cognitive functioning and safety determinants in the work of a train drivers. *Acta Neuropsychologica* 19(2), 279–291 (2021). <https://doi.org/10.5604/01.3001.0014.9958>
 51. Madlenak, R., Masek, J., Madlenakova, L.: An experimental analysis of the driver’s attention during train driving. *Open Eng.* 10(1), 64–73 (2020). <https://doi.org/10.1515/eng-2020-0011>
 52. Suzuki, D., Yamauchi, K., Matsuura, S.: Effective visual behavior of railway drivers for recognition of extraordinary events. *Q. Rep. RTRI* 60, 286–291 (2019). https://doi.org/10.2219/rtriqr.60.4_286
 53. Silversmith, D., et al.: Plug-and-play control of a brain–computer interface through neural map stabilization. *Nat. Biotechnol.* 39(3), 326–335 (2021)
 54. Zeng, Y., Sun, K., Lu, E.: Declaration on the ethics of brain–computer interfaces and augment intelligence. *AI Ethics* 1(3), 209–211 (2021). <https://doi.org/10.1007/s43681-020-00036-x>
 55. Chunsheng Wanga, Hui Yi, Wei Wang, Palaniappan Valliappan. Lesion localization algorithm of high-frequency epileptic signal based on Teager energy operator. Volume 47, 2019, Pages 262-275, ISSN 1746-8094, <https://www.sciencedirect.com/science/article/abs/pii/S1746809418302313>
 56. Mique Saesa, Carel G.M., Meskers, Andreas Daffertshofer, Erwin E.H., van Wegen, Gert Kwakkel. Are early measured resting-state EEG parameters predictive for upper limb motor impairment six months poststroke? Volume 132, Issue 1, 2021, Pages 56-62, ISSN 1388-2457, <https://doi.org/10.1016/j.clinph.2020.09.031>.
 57. C. W. Martin, Ed., *The philosophy of deception*, 1st ed. Oxford University Press on Demand, 2013, pp. 3–11, ISBN :9780195327939.
 58. Chuin Hong Yap, Moi Hoon Yap, Adrian Davison, Connah Kendrick, Jingting Li, Su-Jing Wang, and Ryan Cunningham. 2022. 3D-CNN for Facial Micro- and Macro-expression Spotting on Long Video Sequences using Temporal Oriented Reference Frame. In Proceedings of the 30th ACM International Conference on Multimedia (MM '22). Association for

Computing Machinery, New York, NY, USA, 7016–7020.
<https://doi.org/10.1145/3503161.3551570>.

59. S. P. Teja Reddy, S. Teja Karri, S. R. Dubey and S. Mukherjee, "Spontaneous Facial Micro-Expression Recognition using 3D Spatiotemporal Convolutional Neural Networks," 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-8, doi: 10.1109/IJCNN.2019.8852419.
60. Katia Romero, Elisa Yumi, Sandra Camargo, Fabiano Ferrari. Revisão Sistemática da literatura em engenharia de software teoria e prática. Edição: primeira. Ano de publicação: 2017. Editora: LTC. ISBN: 9788535286410.

Modelos de Combinação de Serviços Text-To-Speech com Locução Convencional para a Oralização de Notícias

Abstrato. O aumento do consumo digital tem sido, em muitos casos, penalizador para as empresas de media, levando à redução de receita e à necessidade de reinvenção dos modelos de negócio. A revolução digital veio introduzir novos formatos de consumo e, neste contexto, os podcasts noticiosos são já uma realidade. Embora a sua existência seja relativamente curta no que à cronologia do jornalismo diz respeito, o aumento no consumo deste formato tornam-no num alvo apetecível para incorporar no jornalismo. Este artigo centra-se, essencialmente, na premissa de facilitar a criação deste formato, tendo em mente a redução de custos e conhecimentos necessários para a sua implementação. Para isso, propõe-se o recurso à Tecnologia de Text-to-Speech (TTS) para a oralização dos textos jornalísticos em português de Portugal. Assim, realizaram-se testes utilizando serviços TTS da Amazon Polly e Google Speech Cloud, com os serviços da Google Speech Cloud Wavenet a obterem resultados superiores por parte de potenciais utilizadores. Foram também criados 3 modelos de podcast com incorporação de locução com voz humana e/ou TTS. O grau de existência de voz humana influenciou positivamente os resultados, com o modelo de voz humana e voz híbrida, a obterem melhores resultados que o modelo exclusivo de voz TTS. Contudo, as diferenças são pouco significativas entre os modelos e os resultados evidenciaram uma aceitação da tecnologia Text-to-Speech no uso de podcasts noticiosos. Verifica-se, no entanto, a necessidade de evolução contínua da tecnologia na ótica de convergência com o discurso humano.

Palavras-chave: Jornalismo, multimédia, automatismo, digital, notícias, áudio, locução, podcast, TTS, oralização

1 Introdução

Esta era marcou a passagem de um jornalismo tradicional de imprensa em papel, para um jornalismo digital e online. Esta alteração do modo de consumo reduziu substancialmente a receita proveniente da venda de exemplares físicos que levou à necessidade de reinvenção dos modelos de negócio que vigoravam na indústria do jornalismo para garantir o financiamento das redações [1]. O esforço necessário para esta adaptação afetou as redações de forma diferente consoante a sua capacidade de investimento na modernização. O propósito deste artigo relaciona-se com a oportunidade que resulta do aparecimento de novas tecnologias que permitem reduzir o custo necessário para a adaptação do jornalismo para novos formatos, nomeadamente os Podcasts Noticiosos. O uso de ferramentas automatizadas de produção de conteúdo, nomeadamente, as ferramentas automatizadas de conversão de texto-para-voz TTS (Text-To-Speech), são já uma realidade e vêm sendo gradualmente introduzidas no campo do jornalismo.

2

Sucintamente, as soluções de TTS permitem a conversão de texto para som, simulando a voz humana por via de algoritmos de fonética que, dado os avanços na tecnologia, se assemelham cada vez mais à voz humana. Num contexto de produção de Podcasts Noticiosos, estas ferramentas permitem a eliminação de várias etapas e de material requerido, no sentido em que não necessitam de um locutor, material de captação e conversão sonora e tratamento de áudio. Esta diminuição do esforço necessário para obter como resultado a oralização de conteúdo noticioso poderá ser um tónico motivador para que esta tecnologia seja adaptada por redações que pretendam a conversão das suas notícias sem prejuízo para a qualidade das suas funções. Assim, importa perceber como é que o público percebe estas novas ferramentas e de que forma um Podcast Noticioso, que incorpore esta tecnologia, é recebido por parte do consumidor final. O termo “Podcast” foi apenas cunhado há 17 anos, sendo ainda recente no que à cronologia do jornalismo diz respeito [2]. A compreensão deste fenómeno e a criação de um modelo de equilíbrio entre a locução convencional e a sintetizada são os princípios que guiam este artigo, com a premissa assente na ótica de um contributo para o jornalismo digital e para a sua evolução.

2 Enquadramento teórico

O ciberjornalismo assumiu um papel primordial enquanto fonte para o consumo noticioso, afigurando-se como a normalidade neste novo quotidiano pela facilidade com que a ele recorremos e pela forma como se adapta aos dispositivos à nossa volta [3]. A digitalização do jornalismo, no que ao consumo diz respeito, parece ter vindo para ficar, como demonstram os dados referentes à Alemanha que, entre o período de 2013-2020, revelam um declínio de 37% no uso de jornais como a fonte principal de notícias, a par de um crescimento de 21%, no mesmo período, das redes sociais enquanto fonte noticiosa [4]. Esta migração também se verificou em Portugal o que levou a uma redução das receitas para os jornais e, conseqüentemente, a um desinvestimento na imprensa portuguesa, com uma redução de 75% no investimento entre 2006 e 2016 [5]. Os media tiveram que adaptar a sua produção de conteúdo por forma a darem resposta aos novos métodos de consumo da audiência, nomeadamente os digitais [6]. Entre os formatos que o conteúdo noticioso pode assumir e que regista um forte crescimento está o formato áudio. Neste processo de adaptação para este formato, as ferramentas de TTS permitem a oralização das notícias sem necessidade de gravação de locução. Com o aparecimento dos Podcasts noticiosos, as ferramentas de TTS poderão ser um auxílio importante na criação de conteúdos audiovisuais complementares à notícia. Os Podcasts, pela sua natureza digital, permitem a faculdade de serem partilhados em interfaces que consintam um acesso constante a este tipo de conteúdos em tempo de real e de forma instantânea. Estes fatores tornaram os podcasts um caso de sucesso com um assinalável crescimento na última década [2]. Plataformas como o Spotify registaram subidas de 24% no número de subscritores no ano de 2020, atribuindo esse facto aos podcasts [7]. No caso da categoria específica de podcasts noticiosos, estes continuam a aumentar a um ritmo bastante assinalável. Só entre janeiro e outubro de 2019, o número de podcasts categorizados como “podcasts de notícias” cresceu 32% [8]. Os Podcasts

Noticiosos foram o 2º género mais escutado no segundo trimestre de 2022 nos Estados Unidos, apenas atrás do género de comédia [9]. O uso de TTS permite a eliminação de várias etapas e de material requerido para uma oralização tradicional de uma notícia, ou para produção de um podcast. Este método poderá ser vantajoso para redações que não possam suportar os custos ou o tempo necessário para a produção deste formato

3 Tecnologia TTS

A tecnologia TTS evoluiu imenso desde o seu aparecimento em meados do século XIX [10]. Atualmente, os sistemas de TTS procuram imitar a voz humana compreendendo não só as suas componentes sonoras, mas também as idiossincrasias e nuances dos diferentes tipos de voz [11]. Com o evoluir desta tecnologia e a sua convergência e aproximação à voz humana, a sua integração num processo de oralização de notícias torna-se mais atrativa para o consumidor. Com esta evolução, a voz sintetizada aproxima-se de uma voz de locução convencional permitindo, por exemplo, a sua inclusão num podcast noticioso. No que diz respeito à escolha da tecnologia TTS a aplicar no contexto de um podcast noticioso, apesar da “naturalidade” parecer um critério óbvio quando se pretende escolher um TTS que se assemelhe a uma voz humana, há um grau de subjetividade inerente à função que se pretende que este TTS desempenhe. Esta subjetividade aparenta ter relevância na hora de determinar qual o melhor TTS para uma determinada função [12] pelo que será necessário escolher a melhor metodologia para a avaliação

3.1 A Avaliação da Tecnologia TTS

Um TTS com a função de assistente virtual poderá ser avaliado de forma diferente de um TTS com a função de leitura [13]. Os testes tradicionais de avaliação de TTS têm por base a avaliação de amostras sonoras. Nestes testes, os participantes avaliam amostras e enunciados não relacionados e livres de contexto. Estes testes têm sido contestados por ignorarem o uso real destes TTS no contexto em que estes são aplicados ou para a funcionalidade a que se propõem [12]. Para a avaliação dos sistemas de voz sintéticas era tradicionalmente utilizado o método de MOS1 - Mean Opinion [14]. O método MOS foi, durante anos, o padrão de avaliação das vozes sintéticas. Contudo com os avanços científicos, parece estar limitado enquanto método de avaliação único [15]. Por forma a perceber quais os aspetos que influenciam a preferência, os autores Joseph Mendelson e Matthew Ayelett, propõem um teste que avalia a naturalidade da voz e a sua expressividade [12]. Para este efeito, o teste utiliza vozes com evidentes diferenças de tom, pretendendo-se que sejam avaliadas as características de naturalidade e expressividade de 1 a 5, consoante o método de MOS. Estas amostras sonoras são utilizadas no mesmo contexto com a reprodução de excertos de texto idênticos. Assim, é assegurado que os participantes interpretam qual a melhor voz para o caso específico que lhes é apresentado e não qual a melhor voz em geral [12]. Outro caso de inovação no estudo e avaliação da voz sintética avalia os TTS em diferentes contextos, declarados à priori, tendo em conta a função a que se destinam, mas também o resultado da interação entre o ouvinte e o TTS no caso de a função exigir uma resposta [13].

4

4 Metodologia

A metodologia utilizada por este estudo pressupõe uma abordagem teórica à questão enfoque, à qual se segue uma fase de design experimental, originando um desenvolvimento de protótipos que ofereçam resposta à problemática que dá o título a este artigo. Três fases foram determinadas: em primeiro lugar, a avaliação de um serviço de TTS, adequado para aplicação em contexto comercial, que permitisse a conversão de conteúdo noticioso em texto para voz, utilizando a tecnologia TTS em Português de Portugal. Após a determinação desse serviço, o segundo objetivo fixou-se na criação de modelos de podcasts noticiosos com diferentes graus de presença de TTS. Por último, os modelos foram avaliados junto de uma audiência de podcasts noticiosos.

4.1 Primeira avaliação

Pretendeu-se, nesta primeira fase, averiguar quais os serviços de TTS que colhiam melhor receptividade por parte do público. Os serviços foram avaliados com recurso à conversão de um excerto de texto, retirado de uma notícia real, para áudio por via de conversão via TTS. Estes excertos foram disponibilizados numa plataforma de inquéritos, de administração online. Foram avaliadas 6 soluções de TTS - 2 serviços em Português de Portugal disponibilizados pela Amazon Polly (Inês Polly (V3) e Cristiano Polly (V5)) e os 4 serviços disponibilizados pela Google Cloud Speech WaveNet (pt-PT-Standard-A Wavenet (V1), pt-PT-Standard-C WaveNet (V2), pt-PT-Standard-B WaveNet (V4) e pt-PT-Standard-D WaveNet (V6)). A avaliação teve em conta as seguintes 5 características: Qualidade Geral (P1), Pronúncia (P2), Agradabilidade (P3), Enquadramento na Tarefa (P4) e uma escolha final da Voz favorita para a função de Oralização de Notícias (VF). As opções de resposta foram colocadas de forma que pudessem ser convertidas para avaliação psicométrica via escala de Likert [16]. A escolha dos 6 TTS teve em conta o estado da arte em estudos comparativos anteriores sobre a tecnologia TTS nomeadamente o estudo comparativo de [17] e o estudo sobre a perceção da qualidade dos TTS [13]. A amostra por conveniência incluiu 53 inquiridos, sendo 37 do género masculino e 16 do feminino. As idades dos participantes encontram-se compreendidas entre os 19 e os 65 anos, com uma média de 30 anos ($M = 30$; $DP = 10.64$).

Table 1. Tabela resumo dos resultados relativos à avaliação dos Serviços TTS

	P1	P2	P3	P4	Média	VF
V1	3.26	3.77	3.04	3.06	3.28	5.70%
V2	3.79	4.32	3.55	3.55	3.8	15.10%
V3	2.47	3.19	2.36	2.49	2.63	7.50%
V4	3.89	4.47	3.75	3.89	4	43.40%
V5	2.23	2.85	2.32	2.21	2.4	1.90%
V6	3.94	4.53	3.83	3.90	4.05	26.40%

As vozes com melhor qualificação nos parâmetros correspondem às vozes que reuniram maior preferência (Voz 4 – 43,4% e Voz 6 26,4%). Embora a Voz 6 tenha obtido uma média superior nos parâmetros individuais, na pergunta final sobre qual a voz preferida, esta foi preterida pela voz 4. As duas vozes com piores resultados pertencem ao serviço da Amazon Polly (Voz 3 e Voz 5.) Ficam evidenciadas, nos resultados, as diferenças entre as duas tecnologias em relação à sua aplicabilidade para a função leitura de notícias, em português de Portugal. Identificadas as vozes preferidas, o estudo prosseguiu com avaliação de modelos de podcasts que integrassem soluções TTS.

4.2 Criação e avaliação dos Modelos de Protótipos

Para avaliar de forma comparativa os modelos, determinou-se a criação de um podcast de voz humana (R), um podcast de voz sintetizada (S) e um podcast que utilizasse tanto a voz humana como a voz sintetizada, denominado “podcast híbrido” (H). Os modelos partilharam textos e uma estrutura de base comum composta por: Jingle introdutório, Música de fundo, Matéria redatorial, Separadores sonoros, Notícias, Jingle final. Seguidamente procedeu-se à análise dos modelos de combinação de serviços text-to-speech com locução convencional. Para se proceder a uma avaliação eficiente dos modelos gerados foi determinada a administração de um inquérito com as seguintes questões: *Em relação ao Podcast “x” qual foi a impressão geral acerca do que ouviu (IG); Em relação à locução do Podcast “x” indique o grau de adequabilidade das vozes (LO); A voz sintetizada poderá ser utilizada no contexto de podcasts? indique o seu grau de concordância;* As opções de resposta a estas três questões foram colocadas de forma a serem expressos em Escala de Likert. Por fim, uma última questão: *Consideraria escutar um podcast de voz sintetizada no futuro?* Posteriormente, foi conduzida uma entrevista semiestruturada individual, para que os participantes pudessem expor as suas opiniões acerca dos modelos escutados. Foi objetivo desta entrevista a identificação das oportunidades de melhoria relativas à Tecnologia TTS no Contexto de Podcasts Noticiosos. No que diz respeito à amostra, participaram 15 indivíduos. Em relação à amostragem esta foi não probabilística por conveniência. Em relação ao sexo dos participantes, 8 (53%) identificam-se como homens e 7 (47%) como mulheres.

Table 2. Fig. 1. Tabela 3. Resumo dos Dados Relativos À Impressão Geral dos Modelos e Adequabilidade das vozes para a Locução (Esquerda); Uso de vozes sintetizadas no contexto de Podcasts Noticiosos (Meio); Perspetiva de escutar um Podcast Noticiosos TTS no Futuro (Direita)

Modelos Podcast	IG	LO	Contexto Podcast		Podcast TTS Futuro	
	Likert		Likert		Sim	5
H	4.1	3.8	Média	3.53	Após melhorias	10
R	4.3	4.1	Mediana	4	Não	0
S	3.9	3.7	D. Padrão	1.13	Total	15

6

No que à Impressão Geral diz respeito observamos que o modelo com Voz Real assume os valores mais altos, com 4,27, seguida do Podcast de Voz Híbrida com 4,07 e por fim o modelo com Voz Sintetizada com 3,93. Em relação à locução o Podcast de Voz Real, obtém o valor médio mais elevado (4,07), seguido do de Voz Híbrida (3,80) e por fim novamente o de Voz Sintetizada (3,67). Em relação à possibilidade de uso de vozes TTS no contexto de Podcasts noticiosos os inquiridos indicaram uma média de concordância de 3,53 em escala de Likert. Aproximadamente 66% das respostas dadas a esta questão (n=10) foram de tendência de concordância positiva. Por fim, analisando os dados da última tabela, analisamos que 5 utilizadores responderam estar dispostos a escutar um podcast TTS no futuro e outros 10 referiram que o fariam após melhorias à tecnologia atual, não se registando qualquer resposta negativa.

5 Conclusão

No que diz respeito às vozes em português de Portugal, naquela que correspondeu à seleção inicial que veio a definir todo o processo subsequente, as vozes 4 e 6 do serviço WaveNet da Google Cloud Speech reuniram a maior preferência (Voz 4 – 43,4% e Voz 6 26,4%) tendo sido também as mais bem pontuadas nos parâmetros individuais. Já as vozes do serviço da Amazon Polly, obtiveram as piores pontuações individuais, com alguma discrepância face aos serviços da WaveNet. No Questionário Final, constatou-se uma forte tendência de aceitação da integração da tecnologia Text-to-Speech. No questionário está subjacente a tendência para aceitar a convergência da tecnologia Text-to-Speech em Voz Real. Esta convergência fica patente nas necessidades de melhoria sublinhadas pelos utilizadores acerca da voz sintetizada. É notória a vontade de que o TTS se aproxime cada vez mais da voz real, sendo que a aceitação das vozes sintetizadas tem, aparentemente, melhores níveis quanto mais se aproxime da voz humana. Os testes de validação dos modelos mostraram precisamente que os modelos com mais presença de voz humana obtêm melhores pontuações. Quer o modelo de voz real, quer o modelo de voz híbrida obtiveram pontuações superiores ao modelo de voz TTS nos parâmetros analisados, sendo que o modelo de Voz Real foi o mais bem pontuado. No entanto, a discrepância entre os valores obtidos por cada modelo é reduzida, com 4,07 para o modelo de Voz Real, seguido da Voz Híbrida com 3,80 e, por fim, a voz sintetizada com 3,67. Nesse sentido, todos os modelos apresentaram uma receptividade bastante positiva. Assim, pode-se especular que o sucesso da voz sintetizada no contexto de podcasts noticiosos ou na oralização de notícias parece depender do seu grau de aproximação à Voz Humana e não de uma voz puramente sintetizada com características diferentes da voz humana. Contudo, dada a reduzida expressão das diferenças de receptividade dos diferentes tipos de vozes, identificam-se oportunidades para utilização alargada das soluções TTS, considerando também os fortes investimentos e melhorias que se verificam nestas soluções.

Referências

- [1] D. Harte, R. Howells, and A. Williams, *Hyperlocal journalism: The decline of*

- local newspapers and the rise of online community news*. Routledge, 2018.
- [2] N. Newman and N. Gallo, “News podcasts and the opportunities for publishers,” 2019.
- [3] S. Allan, *Online news: Journalism and the Internet*. McGraw-Hill Education (UK), 2006.
- [4] N. Newman, with Richard Fletcher, A. Schulz, S. Andi, and R. Kleis Nielsen, “Reuters Institute Digital News Report 2020,” 2020.
- [5] M. Botelho, “A crise dos jornais e do jornalismo. Meios & Publicidade.” 2017.
- [6] M. Stephens, *A history of news*. Oxford University Press, 2007.
- [7] M. Sweney, “Spotify credits podcast popularity for 24% growth in subscribers | Spotify | The Guardian,” Feb. 03, 2021. <https://www.theguardian.com/technology/2021/feb/03/spotify-podcast-popularity-24-percent-growth-subscribers> (accessed Feb. 23, 2021).
- [8] M. Bhattacharjee, “News podcasts grow by 32% as daily news shows become increasingly popular, reports Reuters | What’s New in Publishing | Digital Publishing News,” Dec. 10, 2019. <https://whatsnewinpublishing.com/news-podcasts-grow-by-32-as-daily-news-shows-become-increasingly-popular-reports-reuters/> (accessed Feb. 23, 2021).
- [9] Edison Media, “Comedy Tops the Podcast Genre Chart in the U.S. for Q2 2022 - Edison Research.” <https://www.edisonresearch.com/comedy-tops-the-podcast-genre-chart-in-the-u-s-for-q2-2022/> (accessed Nov. 05, 2022).
- [10] D. H. Klatt, “Review of text-to-speech conversion for English,” *J. Acoust. Soc. Am.*, vol. 82, no. 3, pp. 737–793, 1987.
- [11] S. O. Arik *et al.*, “Deep voice: Real-time neural text-to-speech,” *arXiv Prepr. arXiv1702.07825*, 2017.
- [12] J. Mendelson and M. P. Aylett, “Beyond the Listening Test: An Interactive Approach to TTS Evaluation.,” in *INTERSPEECH*, 2017, pp. 249–253.
- [13] P. Wagner *et al.*, “Speech synthesis evaluation—state-of-the-art assessment and suggestion for a novel research program,” 2019.
- [14] I. Rec, “P. 85. a method for subjective performance assessment of the quality of speech voice output devices,” *Int. Telecommun. Union, Geneva*, 1994.
- [15] T. Hoßfeld, R. Schatz, and S. Egger, “SOS: The MOS is not enough!,” in *2011 third international workshop on quality of multimedia experience*, 2011, pp. 131–136.
- [16] R. Likert, “A technique for the measurement of attitudes.,” *Arch. Psychol.*, 1932.
- [17] J. Cambre, J. Maddock, J. Tsai, and J. Colnago, “Choice of Voices: A Large-Scale Evaluation of Text-to-Speech Voice Quality for Long-Form Content,” vol. 20, Apr. 2020, doi: 10.1145/3313831.3376789.

Field trial of an iTV solution to send personalized notifications

João Encarnação¹ [0000-0002-9735-183X], Ana Velhinho¹ [0000-0001-9978-8317], Simão Bentes¹ [0000-0001-5195-7839], Telmo Silva¹ [0000-0001-9383-7659] and Rita Santos² [0000-0001-9741-6210]

¹Digimedia, Department of Communication and Art, University of Aveiro,
3810-193Aveiro, Portugal

²Digimedia, Águeda School of Technology and Management, University of Aveiro,
3754-909 Aveiro, Portugal
{jrangel29, ana.velhinho, bentes, tsilva, rita.santos}@ua.pt

Abstract. In the context of the proliferation of push notifications on several devices, a mechanism of personalized notifications is proposed to the TV ecosystem to assess its potential. After validating use scenarios through focus groups, a field trial was carried out to test the solution using a managing platform, developed to schedule personalized notifications to users' set-top-boxes (STB). Over different timeframes and STB triggers, several notifications' categories were tested, namely: whether forecast; morning and evening local news with the highest index of credibility; calendar appointments and local events; birthday reminders; medication alerts; wellbeing recommendations; suggestion of food delivery services on defined schedules; and reminders that a regular show is starting in 5 minutes enabling to directly switch the channel. The field trial allowed measuring the acceptance of the proposed solution, regarding the schedules and different categories, as well as gathering suggestions for improvements and other relevant notifications.

Keywords: Push notifications; iTV; Field Trial; Notifications manager prototype

1 Introduction

Push notifications constitute a message mechanism that catches users' attention multiple times per day on their devices, thus becoming a prevalent way of accessing information. Statistics for 2022 indicate that people are more receptive to push notifications if they can choose from which apps they wish to receive them, giving relevance to personalization features and sending schedules, which can increase click rates by 40%[1].

Complementary to the previous literature review within this research, about the potential of TV notifications and related work regarding users' evaluation of prototypes [2], other studies highlight the importance of notifications' temporality. Namely, by identifying the moments when people value the most and dedicate more time to information and news messages on mobile devices [3]. In the television context, notifications are a less explored mechanism, because TV continues to be used

collectively and may raise some privacy issues [4]. Nevertheless, the proliferation of Smart TVs and Hybrid Broadcast Broadband TV (HbbTV) reinforced the Smart Home experience [5] and boosted marketing tools based on notifications to generate and measure engagement, related to viewing behaviours towards apps and streaming platforms. Natively, Amazon Fire TV incorporates Amazon Device Messaging (ADM) that generates analytics for measuring engagement but also admits Android Notifications API for sending local notifications outside the app’s UI of Fire TV [6]. In the case of Netflix, it incorporates a Rapid Event Notification System (RENO), which uses a hybrid push-and-pull communication model across various platforms and devices [7].

In this context, the OverTV project results from a partnership between academia and an IPTV provider with the aim to develop a solution for personalized notifications towards the TV ecosystem, mediated by a web platform [8] developed to generate, schedule, and monitor several types of messages (Figure 1). From the requirements and scenarios most valued by a focus group of potential users, the system includes five thematic categories of notifications: “Info”, “Content”, “Services”, “Calendar”, and “Health”. Hence, the notifications are based on geo-referenced data and users’ preferred content regarding information, entertainment, and services, as well as calendar and health reminders. The same event may generate several notifications depending on each thematic category’s predefined settings (which are also customizable). These parameter settings are related to STB triggers and timings which were tested in a field trial.

This paper presents partial results of the field trial of an iTV solution to send notifications to TV set-top boxes (STBs). Because the TV notifications followed the interface style and remote-control interactions of the IPTV partner, this study is not centred on usability metrics but on understanding if the notifications from several thematic categories and their different schedules were useful in the users’ daily life. The paper is divided into four sections: the introduction that presents the scope and objectives of the solution; the methodology that includes the data collection procedures and the sample; the results section with the main insights from the field trial; and the final considerations with an overview of the study and future steps.

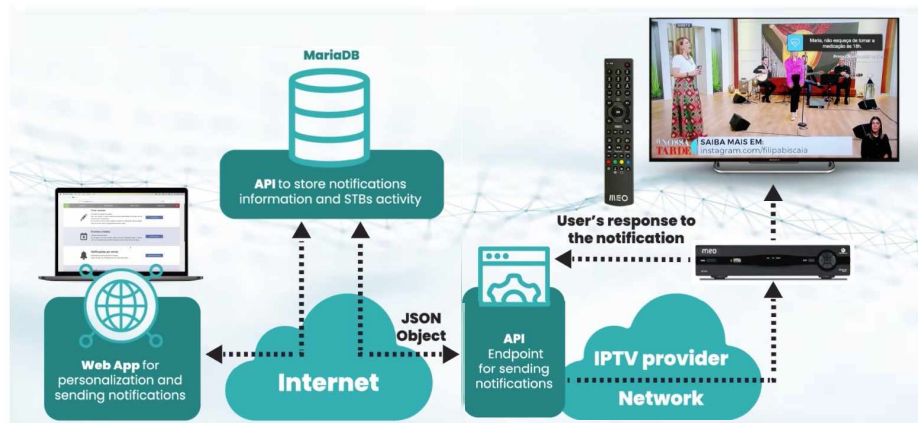


Fig. 1. System architecture

2 Methodology

The main goals of the field trial were to understand how the users react to TV notifications in a shared device and how useful the various categories and schedules of notifications were to their daily lives. After the users' experience during the trials, the aim was also to obtain suggestions for improvements and other typologies of notifications. The sample comprised 25 participants, 15 female (60%) and 10 male (40%), between 23 and 85 years old. Despite the small size of the sample, the criteria to participate were having the STB service of the partner IPTV provider (MEO) and being available to participate in an ongoing study with a qualitative focus. Another important factor for the sample was including diversified household dynamics. Hence, the 25 participants are distributed into 14 houses (each corresponding to an STB) with different compositions of family households: singles, couples, families with children, families with adults and elderly.

Given the wide age range of the sample, some participants expressed difficulties using certain devices. Amongst the elder participants, 7 users mentioned that they face difficulties using the smartphone, whilst the other 2 do not use this device at all. Regarding the tablet, 10 participants do not use it and another 2 have difficulties doing so. Finally, 4 participants expressed difficulties using the computer and 7 do not use it at all. Regarding the TV consumption habits of the sample, 56% of the participants mentioned that they watch TV more than once a day, and 36% said they watch it once a day. Only 1 user mentioned watching TV once a week and another user watches once a month.

The field trial was carried out for 44 days, between the month of July and August 2022. The evaluation protocol was divided into four phases: 1) the recruitment of participants and providing instructions about the proposed solution and the field trial; 2) filling out a pre-test survey to characterize the sample regarding the usage of devices and TV consumption habits, and also to gather users' preferences and information to create personalized notifications; 3) implementation of the field trial in the participants' households, with permanent remote monitoring and technical support when needed; 4) filling a post-trial survey to obtain feedback, namely about the receptivity to the service and the several thematic notifications, as well as the identifications of pain points and suggestions for improvements and other notifications.

3 Results and discussion

The results obtained from the field trial were gathered using the statistics provided by the notification management platform, developed within the research project. It is important to clarify that the notifications were only considered successfully sent if the user's TV STB was turned on. Nevertheless, it was not possible to confirm if and which user of the household viewed the notification. For this purpose, it was used complementarily the data from the post-trial survey which considers the users' perceptions about the number of notifications they viewed and their usefulness by thematic category (Table 1).

Table 1. Comparison by category of rates of sent notifications and of rates of usefulness.

Rates of sent notifications	Category	Highlights of rates of perceived usefulness
<i>Total</i>		
72,5%	<i>All</i>	32% “very useful” 52% “useful” 04% “useless” 12% “very useless”
<i>General notifications</i>		
93,4%	<i>Morning and evening routines</i>	48% “very useful” 36% “useful”
55,5%	<i>Information</i>	40% “useful” (<i>morning news</i>) 40% “did not receive any” (<i>evening news</i>)
<i>Personalized notifications</i>		
84,5%	<i>Services</i>	24% “useful” 40% “did not receive any”
76,3%	<i>Content</i>	32% “very useful” 36% “useful”
53,9%	<i>Health</i>	48% “very useful” (<i>medication</i>)
52,0%	<i>Calendar</i>	32% “very useful” 44% “useful” (<i>local events</i>)

Overall, throughout the field trial, 4465 notifications were created, of which 72,5% were successfully sent (3693). When questioned about how many notifications they viewed, 32% of the users responded “more than 3 per day”, while 8% said they “never saw any notification”. Complementing these statistics, 64% of the participants considered that the number of notifications they saw every day was “adequate”, and another 12% considered it “very adequate”.

Within the proposed system of notifications, it is possible to distinguish two main types: *general notifications* and *personalized notifications* (Table 1). The difference between the two is that *general notifications* are sent to everyone with the same content and schedule, while *personalized notifications* include the name of the user and information directed to that person (according to data and preferences previously provided by the participants through the survey applied before the field trial).

The *general notifications* include two regular *routines* with information about the date, weather forecast and daily news: the *morning routine* (6 am – 12 am), in which 76,3% of the created notifications were successfully sent, and the *evening routine* (7 pm – 12 pm), in which 84,5% of the notifications were also successfully sent. As shown in Table 1, the *routines* were the category with the highest percentage of received notifications (93,4%), in particular the *evening routine* (84,5%). This suggests that the evening period is more likely to have people watching TV or, at least, having the device turned on. In the post-trial survey, 48% of users expressed that the *routines’* notifications were “very useful” and 36% considered them “useful”. However, 40% of the users considered particularly “useful” the daily *news* sent in the

mourning routine, suggesting that this type of content may be more suited to the beginning of the day. Or this value might also have been influenced by 40% of users mentioning that they “did not receive any” *news* in the *evening routine*. This can be explained by several factors: the fact that the *routines* comprise a sequence of two or more *information* notifications (and sometimes *calendar* notifications), and the users might not pay attention to all the queued notifications; the user had the television connected but was not viewing TV or was not paying attention to the TV in the sending schedule; the system may have registered errors.

Analysing Table 1 helps to understand which schedules were more successful and which thematic categories were perceived as more useful, despite their delivery rates. For instance, the percentage of notifications sent in the category of *services* was the highest (84,48%) among all the *personalized notifications*, although only 24% of the users considered it “useful” and 40% said that they “did not receive any” notification of this kind (only 5 households mentioned habits of ordering food so only were created notifications for those STBs). The success of delivery can be due to the schedule of the *services*’ notifications being close to mealtime since they were only used in the field trial for ordering food services. This suggests the potential of these time slots for other categories of notifications.

The *content* category had the second-best rate of notification delivery (76,3%). These notifications are displayed 5 minutes before the user’s favourite shows start, allowing him to directly switch to the channel the program airs. Not every user mentioned having a TV show or content that they regularly watch on the TV, with most of them only mentioning channel preference. Still, considering that the *content* notification is sent shortly before the users’ most watched shows start, to remind them not to miss it, led to 32% of users considering this category “very useful” and another 36% considered it “useful”, indicating a positive receptiveness to this strategy.

Looking at Table 1, the *calendar* category had the lowest rate of sent notifications (52,03%). This might have happened because the sending parameters for the *calendar* category were to send the notification when the user was zapping through channels and, complementary, also sent it on the day and previous day of the event, integrated into the *morning and evening routines*. Being the *routines*’ schedule the most successful in terms of delivery rates is possible that the zapping trigger might have negatively impacted this value. However, despite the lower rate of delivery, 32% of users found the *calendar* notifications to be “very useful”. Also, 44% considered “useful” notifications that suggested *local events*, which is a type of notification also integrated into the *calendar* category. Furthermore, 23,5% of the users mentioned they “attend to some of the suggested events” and 35,3% said they “did not attend but had an interest in attending”, which shows that this type of notification can have a real impact on people’s lives. Additionally, even users who mentioned they “did not receive any” notification about *local events*, mentioned in the opened suggestions that they found this possibility interesting.

Finally, the *health* category also showed a low percentage of successfully sent notifications. But, in terms of perceived usefulness, the *medication* notification (integrated into the *health* category), presented the highest percentage of being considered “very useful” (48%), along with the *routines* (48%). The low rates of delivery of the *health* category, namely the *medication* notifications corresponding to

intake schedules, may be explained because the users were not watching TV during that period. This fact suggests the need for redundant notifications in personal mobile devices that people likely carry with them indoors and outdoors, namely the smartphone and the tablet. Nevertheless, when asked on which devices they would prefer to receive notifications, only 16% of the participants chose the option of integration of connected devices (e.g., TV, mobile devices, and smart assistants), while 44% preferred exclusively mobile devices and 56% preferred the shared television.

In the post-trial survey, 84% of the users showed a positive reaction to the notifications on the TV, from which 32% found them “very useful” and 52% found them “useful”. Finally, if the system was revised according to their suggestions, 80% of the users showed interest in adopting this service in the future (20% were “very interested” and 60% were “interested”).

4 Final considerations

To evaluate a solution of notification on the TV a field trial was carried out with a sample of 25 participants during a period of 44 days. Given the extent of the results, only a portion of the main insights is presented in this paper, which does not include an in-depth analysis of the suggestions for other features provided by the participants and the comparison of the results according to age and family household dynamics. Overall, there was a positive response to TV notifications in terms of the usefulness of the messages (85%) and a manifest willingness to use such a service in the future (80%). Despite the delivery rates (conditioned to having the TV STB turned on) registered during the field trial, which is associated with different schedules per category, the participants indicated the following preference of time slots to receive notification on the TV: after 7 pm (52%); when turning on the TV in the morning (36%); at lunchtime (24%); or after 10 pm (20%). Elderly participants also indicated the afternoon period (20%) and the morning period (8%).

Even though the best rates of sent notifications are the *morning and evening routines* (93,4%) and *food delivery services* (84,48%), the highest percentages of perceived usefulness were assigned by users to *medication notification* (48%) integrated into the *health* category, along with *notifications about local events* (44%) of the *calendar* category. Both the *health* (53,9%) and *calendar* (52,03%) categories were the ones with the lowest percentages of notifications successfully delivered, which may indicate problems with the timeslots since notifications are only sent the TV STB is turned on, and that does not necessarily mean that the user is paying attention to the TV screen or that other household member may have reacted to their notification since the television is still a shared device. But because both cases of medication intake and appointments have strict timings to be reminded, the way to avoid missing notifications when people are not watching TV or are not at home is a redundant sending of notifications to mobile devices. Nevertheless, 56% of the participants indicated that they preferred receiving notifications on the home-shared television, whilst only 16% chose the option of receiving notifications on several devices at the same time. Because of the current prevalence and familiarity, 44% of

the participants also indicated that they prefer to receive notifications exclusively on their mobile devices. Finally, the field trial results suggested that user-dependent sending parameters, such as zapping, may not be a good option. Since the only category that uses this mechanism is the *calendar* presented the lowest sending rate (52%).

Overall, the field trial validated all the categories included in the prototype, although with the need of fine-tuning the sending schedules and delivery devices to increase the rate of successfully sent notifications. A second field trial will be implemented to allow corroboration of the impact of optimizing delivery times. The other suggested improvements, regarding adding new categories of notifications, allowing synchronization with mobile devices, and providing users' autonomy to schedule notifications through the management platform, will be addressed in future iterations of the prototype.

Acknowledgements

Altice Labs@UA, a research group resulting from the partnership between Altice Portugal and the University of Aveiro, funded this research.

References

1. Business of Apps, Push Notifications Statistics, <https://www.businessofapps.com/marketplace/push-notifications/research/push-notifications-statistics/>, last accessed 2022/11/30.
2. Velhinho, A., Camargo, J., Silva, T., Santos, R.: The Importance of Personalization and Household Dynamics for Notifications in the TV Ecosystem. Applications and Usability of Interactive TV. jAUTI 2021. Communications in Computer and Information Science, vol 1597, 3–19. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-22210-8_1
3. Wheatley, D., Ferrer-Conill, R.: The Temporal Nature of Mobile Push Notification Alerts: A Study of European News Outlets' Dissemination Patterns. Digit. Journal. 9, 694–714 (2021). <https://doi.org/10.1080/21670811.2020.1799425>.
4. Silva, L.A., Leithardt, V.R., Rolim, C.O., González, G. V, Geyer, C.F.R., Silva, J.S.: PRISER: Managing Notification in Multiples Devices with Data Privacy Support, (2019). <https://doi.org/10.3390/s19143098>.
5. Gavrilă, C., Popescu, V., Fadda, M., Anedda, M., Murrioni, M.: On the Suitability of HbbTV for Unified Smart Home Experience. IEEE Trans. Broadcast. 67, 253–262 (2021). <https://doi.org/10.1109/TBC.2020.2977539>.
6. Amazon, Amazon Device Messaging (ADM) and local notifications on Fire TV, <https://developer.amazon.com/apps-and-games/blogs/2021/07/tutorial-amazon-device-messaging-on-fire-tv>, last accessed 2022/11/30.
7. Netflix TechBlog, Rapid Event Notification System at Netflix, <https://netflixtechblog.com/rapid-event-notification-system-at-netflix-6deb1d2b57d1>, last accessed 2022/11/30.
8. Encarnação, J., Velhinho, A., Bentes, S., Silva, T., Santos, R., Salgado, S., Afonso, F.: A management system to personalize notifications in the TV ecosystem. Procedia Computer Science, Volume 219, 674–679 (2023). <https://doi.org/10.1016/j.procs.2023.01.338>

e-inclusion |

Librería de ejercicios en Smart TV para adultos mayores con riesgo de caídas

Magdalena Rosado¹⁻²[0000-0003-2519-4780], María José Abásolo¹[0000-0003-4441-3264],

Telmo Silva³[0000-0001-9383-7659], Stalin Jurado²[0000-0002-1163-1030]

¹Facultad de Informática Universidad Nacional de la Plata (UNLP), Argentina, ²Facultad de Ciencias Médicas Universidad Católica de Santiago de Guayaquil, Ecuador, ³Universidad de Aveiro, Portugal

maria.rosadoa@info.unlp.edu.ar, mjabasolo@lidi.info.unlp.edu.ar, tsilva@ua.pt, stalin.jurado@cu.ucsg.edu.ec

Resumen. Las caídas son un síndrome geriátrico multifactorial y se consideran un problema de salud pública, es la segunda causa de muerte por accidentes o lesiones accidentales, por lo que se busca promover soluciones tecnológicas interactivas que ayuden a reducir el número de personas que presentan riesgos de caídas. Este artículo presenta una plataforma de emisión de contenido en línea en un Smart TV usando Plex, el cual se puede acceder a una librería de vídeos, de forma remota desde cualquier parte del mundo con conexión a internet. La librería contiene rutinas de ejercicios de flexibilidad, fortalecimiento muscular y equilibrio, que son recomendados como parte de un tratamiento de rehabilitación para adultos mayores, a los cuales se les realiza un diagnóstico previo que permite conocer el riesgo de caída que presenta.

Palabras claves: Plex – Smart TV - Ejercicios – Adulto Mayor- Vídeos

1. Introducción

La actividad física para los adultos mayores es esencial para mantener una buena salud. Los fisioterapeutas geriátricos entienden que en esta etapa de la vida empiezan a aparecer ciertas limitaciones motoras o funcionales, y utilizan sus técnicas terapéuticas para ralentizar estos procesos y mejorar la calidad de vida y que en algunos casos pueden llevarse a cabo en el domicilio del paciente.

Los fisioterapeutas adaptan los programas de ejercicios para corregir problemas específicos que aumentan el riesgo de caídas [1].

En la actualidad, existen varios programas de ejercicios de caídas para que los adultos mayores los realicen de forma individual o en grupo sin comprometer su seguridad, pero rara vez se ejecutan; esto se debe al alto costo del transporte al entorno clínico donde se realizan los ejercicios grupales diarios o semanales [2].

Las soluciones tecnológicas podrían proporcionar a los profesionales de la salud una retroalimentación cuantitativa sobre la ejecución de los ejercicios y la progresión de los pacientes, cuando se someten a programas de prevención de caídas [3]. Una tendencia que toma cada vez más fuerza, aquí y en otros países: poner las TIC al servicio de las personas mayores para que puedan disfrutar de vidas autónomas en sus propios hogares [4]. Así pues, el adulto mayor forma parte de una sociedad con cambios vertiginosos, que imprimen un ritmo de vida muy exigente y la constante búsqueda de recursos tecnológicos que conllevan a tener una experiencia televisiva diferente.

En nuestro trabajo previo [5] se presenta el estudio para elaborar una aplicación con contenidos interactivos en Televisión Digital Interactiva (TVDi) para reducir las alteraciones de la marcha en el adulto mayor. Se abordan las consideraciones para la evaluación de la condición funcional a través de un test, el desarrollo de planes de ejercicio localizados, el proceso a considerar para el diseño funcional que tendrá la aplicación interactiva y la planificación de la evaluación del estudio.

El propósito de este estudio es presentar una aplicación que permita acceder a un programa instalado en un servidor NAS a través del cliente Plex en un Smart TV. Esta aplicación incluye una colección de videos de ejercicios diseñados específicamente para rehabilitar y prevenir caídas en personas mayores, permitiéndoles hacerlo desde la comodidad de sus hogares.

El resto de este documento está organizado de la siguiente manera: la sección 2 describe el diseño del sistema, la sección 3 describe cómo funciona el Smart TV y Plex Media Player, la sección 4 describe la aplicación de la herramienta para la prevención de caídas y la sección 5 resume las conclusiones, las principales contribuciones para el trabajo futuro.

2. Diseño del sistema

La investigación desarrollada se basa en diferentes componentes de hardware y software que se combinan para formar un producto interactivo que motiva a los adultos mayores a continuar con un tratamiento de rehabilitación utilizando los resultados de un diagnóstico de riesgo de caída. Estos componentes incluyen el servidor NAS, la aplicación PLEX, Usuarios que interactúan con el Smart TV (Fig.1).

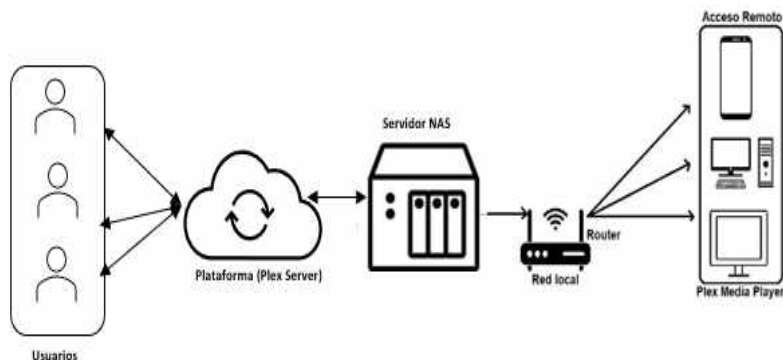


Fig.1 Componentes del sistema

2.1. El Servidor NAS

El servidor NAS (Network Attached Storage) es la solución de almacenamiento para la red que permitirá transmitir los videos de ejercicios almacenados en ese servidor a otros dispositivos clientes.

2.2. Plex Media Server

Plex Server¹ es una aplicación instalada en un servidor o computadora que actúa como un centro de administración y distribución de contenido multimedia. Una vez instalado, Plex Server utiliza el NAS como almacenamiento centralizado para archivos multimedia. Estos son importados y organizados por el NAS, e indexados y catalogados por el administrador en el servidor Plex para crear bibliotecas, carpetas o directorios por tipo de diagnóstico según el riesgo de un caída: leve, moderado y grave (Fig. 2).

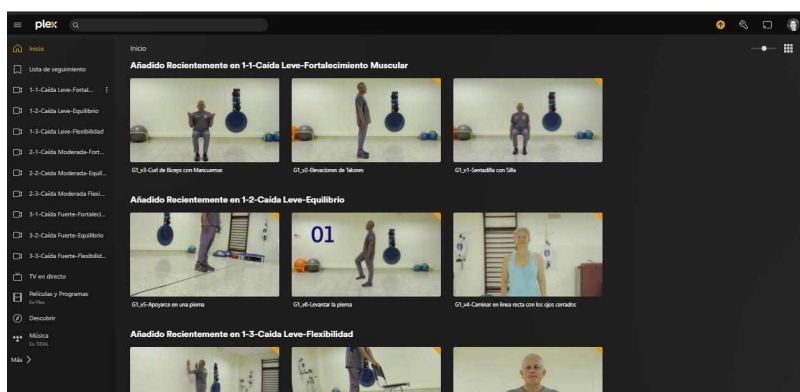


Fig.2. Biblioteca de contenidos de ejercicios

¹ <https://www.plex.tv>

2.3. La aplicación Plex cliente y Smart TV

Plex es una aplicación que se instala en los dispositivos clientes (televisores inteligentes, teléfonos inteligentes, tabletas, computadoras), proporcionando una interfaz intuitiva y amigable que le permite a los usuarios acceder y reproducir archivos multimedia (películas, fotos y escuchar música) en su servidor Plex.

El usuario utiliza un Smart TV o televisor moderno con acceso a internet integrado, un sistema operativo unificado y un procesador similar a una computadora que le permite acceder a productos multimedia interactivos (Fig.3), así como ver videos en línea a través de una conexión de red, ya sea Wi-Fi inalámbrico con cable.

Los dispositivos cliente, tales como Smart TV, se conectan a Plex Server a través de la aplicación Plex, permitiendo acceder y reproducir el contenido multimedia almacenado en el NAS de forma remota. Hay fabricantes de televisores ya incluyen Plex de forma preinstalada en sus dispositivos. En el caso de los televisores que utilizan Android TV como sistema operativo, se puede acceder a Plex a través de la tienda de aplicaciones.



Fig.3. Aplicaciones en Smart TV

3. Aplicación de la herramienta

Para que el usuario (adulto mayor) pueda tener acceso a un contenido multimedia, se configura los permisos de acceso por usuario, permitiendo que el administrador (terapeuta) registre al adulto mayor con su propia cuenta para acceder a la videoteca de ejercicios (Fig.4). Además, el contenido del video tiene una progresión durante la ejecución de los ejercicios; cambiando la velocidad y el tipo de ejercicio.

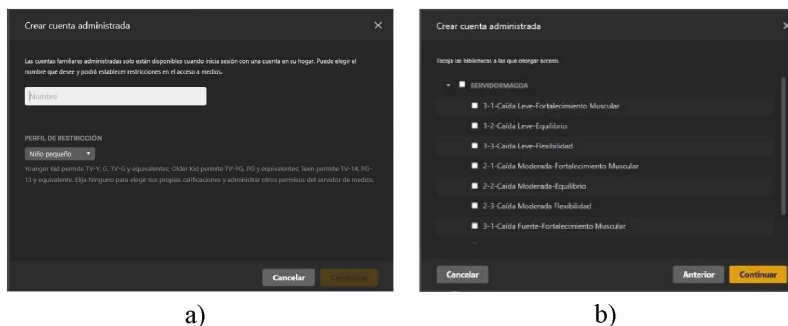


Fig.4. Creación y configuración de accesos a usuarios

Al abrir el reproductor multimedia Plex en el Smart TV, se accede a la interfaz de usuario de Plex, que muestra una biblioteca de ejercicios (fig. 5), permitiendo explorar diferentes categorías, como ejercicios de flexibilidad, fortalecimiento muscular y equilibrio.

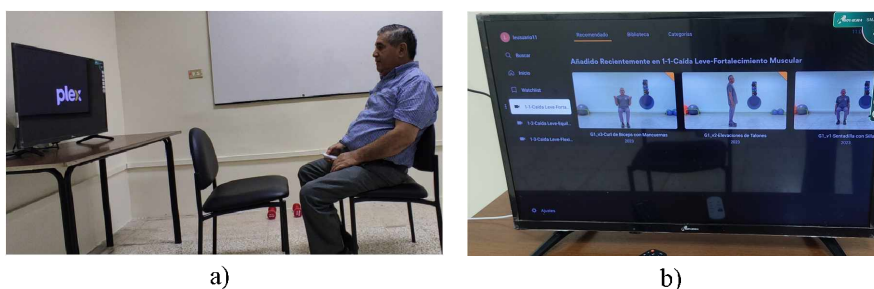


Fig. 5. Sistema de entrenamiento multimedia

El Plex es una herramienta que puede ser de gran utilidad para las personas mayores que necesitan seguir un programa de rehabilitación y estabilizar su marcha después de recibir una evaluación condicional (Fig. 6). A través de un Smart TV conectado a internet, pueden instalar el Plex y acceder a una biblioteca de ejercicios personalizados adaptados a las necesidades individuales.

Dentro de esta biblioteca, se encuentran una serie de videos que explican detalladamente cómo realizar los ejercicios, lo que permite a las personas mayores ejercitarse de manera adecuada y seguir su programa de rehabilitación desde la comodidad de su hogar. Además los videos pueden ser reproducidos tantas veces como sea necesario, lo que les brinda la oportunidad de practicar los movimientos y técnicas específicas hasta que se sientan cómodos con ellos.

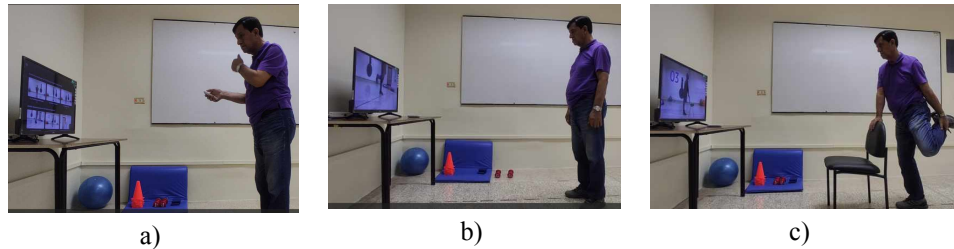


Fig. 6. Experiencia de Smart TV habilitada para PLEX

Asimismo, Plex ofrece la función de marcas de tiempo, que resulta especialmente útil si los videos son largos. Esto significa que las personas mayores pueden detener la reproducción en cualquier momento y luego retomarla desde exactamente el mismo punto donde lo dejaron, brindando la flexibilidad de ajustar su rutina de ejercicios según sus necesidades y capacidad, sin tener que repetir todo el video desde el principio. Para facilitar la navegación por los menús y opciones de la aplicación en el televisor, se utiliza el control remoto para Chromecast. Este mando a distancia permite a los adultos mayores moverse por la interfaz del PLEX de manera más amigable y sencilla, sin la necesidad de utilizar dispositivos adicionales o complicadas configuraciones.

Adicionalmente, el administrador puede ver lo que los usuarios están observando en tiempo real, así como ver el historial de lo que han visualizado, dentro del tablero de actividad en el servidor.

A continuación se presenta el proceso esquemático (Fig. 7) paso a paso para que un adulto mayor pueda encender el televisor y acceder a la aplicación PLEX, con bibliotecas de videos de ejercicios:

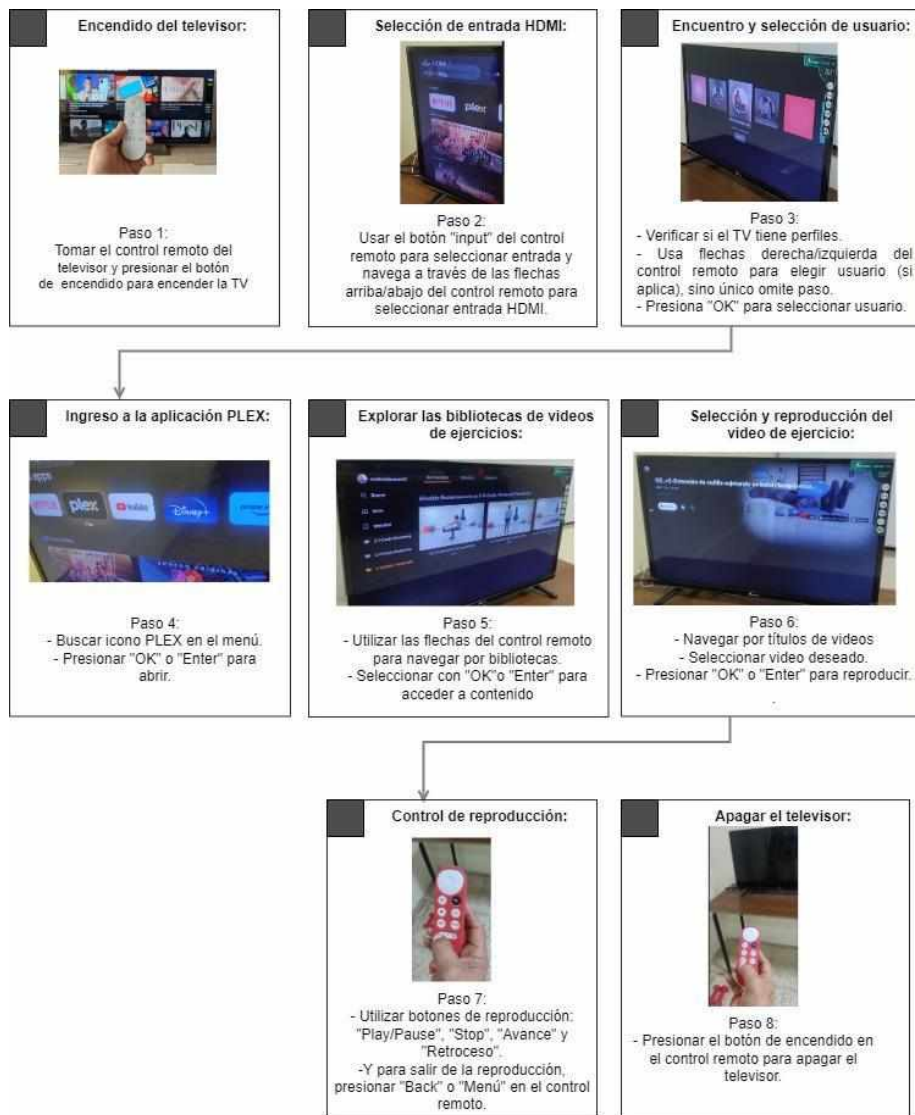


Fig. 7. Sistema de Librería de Ejercicios de Riesgo de Caída en PLEX

4. Conclusión

En esta investigación dedicada a mejorar la marcha en adultos mayores, se destaca la utilización de Plex. Mientras que aplicaciones como YouTube se centran en ofrecer contenido público y compartido en línea basado en recomendaciones y contenido generado por usuarios, Plex ofrece una experiencia personalizada a través de un contenido multimedia para incorporar rutinas de ejercicio físico mientras se disfruta

de la televisión. Esta herramienta resulta especialmente beneficiosa para los adultos mayores que no pueden asistir a un centro terapéutico, permitiéndoles continuar con su tratamiento de rehabilitación y reducir el riesgo de caídas. Para comprobar si el adulto mayor está realizando adecuadamente el ejercicio, en trabajos futuros nos centraremos en la evaluación con herramientas convencionales que permitan evaluar el progreso de los pacientes en términos de composición corporal, fuerza muscular y análisis de la marcha.

Bibliografía

1. Laurence Z. Rubenstein: Caídas en las personas mayores - Geriatría, Manual MSD versión para profesionales(2021).<https://www.msdmanuals.com/es-es/professional/geriatr%C3%ADa/ca%C3%ADdas-en-las-personas-mayores/ca%C3%ADdas-en-las-personas-mayores>
2. Ferreira, B.N., Guimarães, V. y Ferreira, H. S. :Smartphone based fall prevention exercises, en 2013 IEEE 15th International Conference on e- Health Networking, Applications and Services, 643-647 (2013). doi: 10.1109/HealthCom.2013.6720755.
3. Silva, J., Moreira, D., Madureira, J., Pereira, E., Dias, A., y Sousa, I.:A Technological Solution for Supporting Fall Prevention Exercises at the Physiotherapy Clinic, en 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 1-6 (2018). doi: 10.1109/MeMeA.2018.8438811.
4. Rosado M., Abásolo M.J., Silva T. (2020) ICT Oriented to the Elderly and Their Active Aging: A Systematic Review. 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, pp. 134–155 Springer, Cham. ISBN 978-3-030-56573-2 ISSN: 1865-0929 https://doi.org/10.1007/978-3-030-56574-9_9
5. Rosado, M., Abásolo, M. J. y Silva, T. (2021). IDTV Application to Promote the Gait of the Elderly. 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. Springer, Cham. ISSN 18650929 https://doi.org/10.1007/978-3-030-81996-5_10

Las TIC orientadas a las personas mayores con demencia temprana de Alzheimer

Ana Camacho ¹[0000-0001-6689-8667], María José Abasolo ²[0000-0003-4441-3264]
Rita Oliveira ³[0000-0001-6041-9469]

¹ Facultad de Informática de la Universidad Nacional de la Plata (UNLP),
Argentina-Facultad de Ingeniería-Universidad Católica de Santiago de Guayaquil (UCSG),
Ecuador

² Facultad de Informática de la Universidad Nacional de la Plata (UNLP), Argentina

³ DigiMedia, Departamento de Comunicación y Arte, Universidad de Aveiro

ana.camachoc@info.unlp.edu.ar, mjabasolo@lidi.info.edu.ar, ritaoliveira@ua.pt

Abstract. El presente análisis tiene como objetivo indagar metodologías para el mejoramiento de la demencia temprana tipo Alzheimer a través del acompañamiento o asistencia de la tecnología. Se presenta una revisión sistemática de literatura de tipo mixto, transversal entre los años 2017 a 2022, enfocada a la verificación de la existencia de las TIC para adultos mayores entre los 65 y 85 años y su uso efectivo para ayuda en la pérdida de memoria. Los resultados de la revisión revelan que la tecnología para temas de Alzheimer en adultos mayores se direcciona mayoritariamente a la práctica de asistencia y monitoreo, pues justamente es lo que se evidencia como necesidad de las personas con dificultades cognitivas. De igual manera consideran que una estimulación cognitiva mejora la memoria, siendo las soluciones encaminadas a recordar aspectos cotidianos como citas médicas, realizar tareas o tomar medicamentos.

Keywords: Adulto Mayor, Alzheimer, TIC, Tecnología, Televisión Digital Interactiva.

1 Introducción

La enfermedad denominada Alzheimer es una afección crónica neurodegenerativa cuyo rasgo inicial es la disminución de memoria de manera progresiva lo que va dificultando las destrezas para el aprendizaje y la capacidad de realizar actividades diarias en una persona mayor.

La Organización Mundial de la Salud [OMS] en su informe de marzo 2023 sobre demencia indica que los casos diagnosticados con Alzheimer a nivel mundial se encuentran sobre el 60% y 70%, es decir que afecta alrededor de 55 millones de personas y por año obtienen registros de 10 millones de nuevos casos, reportando

así para el año 2030 un incremento de 82 millones de adultos mayores y 152 millones para el año 2050.

Bajo los avances de las Tecnologías de la Información y Comunicación (TIC), éstas pueden facilitar la atención a personas con problemas de memoria. Surge la necesidad de investigar acerca de tecnologías efectivas para personas con Detección Temprana de Alzheimer (DTA), con miras a proporcionar calidad de vida a las personas mayores, así como, ayudar a los cuidadores que mayoritariamente son sus propios familiares sin experticia en el ámbito de ejercicios o métodos preventivos para la pérdida de memoria, logrando capacitarlos en este ámbito con tecnología acorde a los tiempos actuales que facilita la convivencia diaria.

Bustos (2019) destaca que la televisión se posiciona como un medio de comunicación que otorga a las personas mayores entrevistadas canales de interacción para social como sustituto de las interacciones comunicativas. Este resultado invita a reflexionar sobre dicha relación y a colaborar desde el ámbito científico y social para ayudar al sector más mayor de la sociedad.

La Televisión Digital Interactiva (TVDi) actual permite el desarrollo de nuevas plataformas tecnológicas para personas con deterioro cognitivo, especialmente aquellas que padecen la enfermedad de Alzheimer. Estas aplicaciones pueden permitir la estimulación cognitiva que necesitan los adultos mayores.

El presente artículo proporciona la revisión sistemática de literatura sobre la existencia de avances en tecnología para adultos mayores con pérdida de memoria, y en particular desarrollos para la TVDi.

En resto del artículo se organiza de la siguiente forma: La sección 2 describe la metodología aplicada a esta revisión sistemática, la sección 3 muestra los resultados destacando los principales de cada estudio y finalmente la sección 4 presenta conclusiones y trabajo futuro.

2 Metodología

Una revisión sistemática se sostiene por un proceso de investigación cuyo objetivo es obtener, evaluar e interpretar de forma metódica, clara, precisa y rigurosa toda la información que va relacionada con una pregunta de investigación o disciplina. Bárbara Kitchenham (2009) precisa seis pasos fundamentales para la construcción de una exploración veraz de literatura: (i) Formulación de las preguntas que se desea responder respecto a un tema específico, (ii) Búsqueda de las fuentes, (iii) Selección de los estudios a través de los criterios de inclusión y exclusión; (iv) Extracción de información; (v) Presentación de los resultados; (vi) Discusión.

2.1 Preguntas de Investigación

Se plantearon las siguientes preguntas a investigar en relación con las TIC utilizadas para ayudar con el deterioro cognitivo del adulto mayor con DTA:

PI1. ¿Cuáles son las TIC utilizadas?

PI2. ¿Qué propósitos han perseguido las investigaciones encontradas?

PI3. ¿Qué países-continentes se han involucrado en las investigaciones analizadas y en qué fase de desarrollo se encuentran sus desarrollos?

PI4. ¿Cómo se utiliza la TVDi para ayudar a los adultos mayores en el ralentizado del Alzheimer?

2.2 Búsqueda de Fuentes

Se realizó una búsqueda bibliográfica en importantes bases de datos científicas tales como: Scopus, IEEE y Springer. La búsqueda inició desde mayo de 2022, siendo elegidos aquellos artículos íntegros, disponibles y publicados desde el 2017 hasta inclusive el primer semestre del año 2022.

Para una búsqueda asertiva se manejaron descriptores en Ciencias de la Salud (DeCS) presentes tanto en el título como en el resumen de los artículos, siendo la cadena de búsqueda: (Technology AND Elderly AND Alzheimer); y para la nueva tecnología de interés los descriptores (Interactive AND Televisión AND Alzheimer). Se realizó una búsqueda con las palabras claves presentadas en el párrafo anterior dando como resultado un total de 261 artículos (tabla 1).

2.3 Criterios de Inclusión y Exclusión

Se aplicaron los criterios de inclusión y exclusión en 3 fases (tabla 1):

- i. Se eliminaron artículos que mantenían duplicidad en las librerías IEEE y Scopus, de igual manera se eliminó a las que provenían de la misma investigación conservando el que tenía más información. Luego de aplicar estos criterios de exclusión quedaron 94 artículos.
- ii. Se incluyeron artículos relacionados con el estado de desarrollo en el cual se encontraban los hallazgos más relevantes sobre el entorno del adulto mayor y la pérdida de memoria, siendo estos de tipo prototipo, aplicación implementada y de factibilidad, se eliminaron los que no tenían relación con la tecnología o los sistemas informáticos. De igual manera se consideró que la información sea para personas adultas mayores a 65 años. En este caso el número de artículos se redujo a 64.
- iii. Como fase final se trabajó un segundo proceso de exclusión que permitió descartar artículos dirigidos a adultos mayores sanos con inserción de tecnología médica propiamente, estudio de cuidados de pacientes con Alzheimer con herramientas tradicionales, temas de motricidad y el envejecimiento activo, que no son parte de la investigación, por lo cual quedaron 43 artículos que serán utilizados para el análisis de acuerdo con las preguntas planteadas.

Tabla 1. Resultados de la búsqueda y aplicación de criterios inclusión-exclusión

Cadena	Fuente	Total	No duplicado	Criterios inclusión	Criterios exclusión
Technology AND Alzheimer	IEEE	65	64	16	13
AND Elderly	Scopus	20	18	14	7
	Springer	79	79	21	14
Interactive AND Television	IEEE	14	14	2	1
AND Alzheimer	Scopus	35	32	5	3
	Springer	48	48	6	5
Total		261	94	64	43

2.4 Extracción de la Información

Las publicaciones seleccionadas, resultado de aplicar la cadena de búsqueda sobre tecnología y Alzheimer, se clasifican de acuerdo con las siguientes categorías surgidas de las preguntas de investigación.

- Tipo de aplicación:
 - Análisis: análisis de datos en experiencias para promover la inclusión social y el bienestar de adultos mayores (Tabla 2).
 - Lúdica: soluciones enfocadas a ayuda para el adulto mayor a través de juegos (Tabla 3).
 - Compañía: soluciones enfocadas al acompañamiento a personas con deterioro cognitivo (Tabla 3).
 - Asistencia: soluciones enfocadas en ayudar al adulto mayor con sus actividades diarias (Tabla 4).
 - Monitoreo: soluciones enfocadas en supervisar al adulto mayor en procesos diarios (Tabla 4).
- País donde se realizó la publicación;
- Tecnología aplicada;
- Estado del desarrollo: Prototipo (PROTO), Estudio de Factibilidad (FACT) y Aplicación Implementada (APP).

En la Tabla 5 se analizan los desarrollos relacionados con la Televisión Digital Interactiva (TVDi) orientada a personas con deterioro cognitivo.

Tabla 2. Categorización de las publicaciones clasificadas de tipo Análisis

Ref.	Descripción	País	TIC	Estado
[3]	Detección síntomas de Alzheimer	EE.UU.	Móvil/ sensores	APP
[4]	Aceptación de plataforma tecnológica	Grecia	PC/ internet	FACT
[8]	Sistema autónomo colaborativo para adulto mayor	No específica	PC/ sensores	PROTO
[10]	Prácticas de usabilidad de tecnología	Madrid	Móvil	FACT
[13]	Tecnología de apoyo para adaptación	Nueva Zelanda	Móvil	FACT
[17]	Análisis de expresión genética	No específica	PC/ Python	FACT
[19]	Algoritmo para detección de comportamiento	No específica	PC	PROTO
[28]	Interfaz con reconocimiento de voz	Brasil	Móvil/ RA	PROTO
[55]	Clasificación de redes de asistencia a adulto mayor	No específica	PC	FACT
[59]	Diagnóstico de deterioro cognitivo	No específica	PC/ software Magid	FACT

Tabla 3. Categorización de las publicaciones clasificadas por tipo Lúdica y de Compañía

Ref.	Descripción	País	TIC	Estado
Lúdica				
[26]	Juego en Realidad Virtual para adulto mayor	No específica	Móvil/ RV	FACT
[52]	Percepción de usabilidad en videojuegos para adultos mayores	México	PC/ vídeo- juegos	FACT
[60]	Aplicación MovableHealth para mejora cognitiva	Grecia	Móvil/ sensores	FACT
Compañía				
[38]	Acompañamiento gerontecnológico: mascotas virtuales para demencia	EE.UU.	Móvil/ App	FACT

Tabla 4. Categorización de las publicaciones clasificadas de tipo Asistencia y Monitoreo

Ref.	Descripción	País	TIC	Estado
Asistencia				
[27]	Uso de internet por adulto mayor	Alemania	Móvil/ internet	PROTO
[31]	Aplicación PostCard memories	Canadá	Tablet/ RV	PROTO
[11]	Interacción computacional en adulto mayor	Francia	PC	PROTO
[62]	Curso de computación intergeneracional	USA	PC	PROTO
Monitoreo				
[5]	Sistema de monitorización remota para adulto	Portugal	Móvil/ sensores	APP
[18]	Diálogo en Tablet entre cuidador y adulto mayor	Oakland	Tablet	PROTO
[20]	CAL: Hogar inteligente para monitoreo cognitivo	EE.UU.	Móvil/ sensores	PROTO
[45]	Estrategias de I+D+i en TIC para demencia	Canadá	PC	PROTO
[53]	Cuidado robótico para autonomía de adulto mayor	No específica	PC/ sensores	FACT
Asistencia y Monitoreo				
[2]	Sistema biométrico inteligente para Alzheimer	Iraq	Móvil/ sensores	PROTO
[6]	Asist. médica para adulto mayor con MCU ARM	India	Móvil/ sensores	PROTO
[7]	Bastón inteligente para pacientes con Alzheimer	No específica	Dispositivo sensores	PROTO
[15]	Software de realidad virtual para rehabilitación	No específica	Tablet/RV	APP
[22]	Sistema inteligente para reducir limitantes	UK	Móvil/ App	PROTO
[25]	Sistema de entrenamiento para ayuda cognitiva	No específica	Móvil RA	PROTO
[33]	Dispositivo para preservación de recuerdos	No específica	PC/ sensores	PROTO
[34]	Sistema interactivo de localización en tiempo real	Taiwán	PC/ internet	PROTO
[35]	Interfaz interactiva de rastreo de adulto mayor	China	PC/ Python	PROTO
[42]	Sistema de VR para personas con demencia	India	Móvil/ RV	PROTO
[48]	Sistema automático de monitoreo alimenticio para personas con Alzheimer	USA	PC	PROTO

Tabla 5. Caracterización actual de investigaciones sobre la TVDi

Ref.	Descripción	País	TIC	Estado
Análisis				
[21]	Sistema interactivo de TVDi para el entrenamiento cognitivo de pacientes con Alzheimer.	España	Smart TV	APP
[23]	TVDi y sus aplicaciones educativas	México	Smart TV	FACT
[41]	TVDi como herramienta para la estimulación cognitiva en personas con Alzheimer	España	PC	FACT
[58]	Diseño para TVDi social: mejorar la experiencia compartida de los sistemas de atención domiciliaria	España	Smart TV	FACT
[61]	Desarrollo de un episodio piloto de televisión de estimulación cognitiva para personas mayores con demencia	Reino Unido	Smart TV	PROTO
Lúdica				
[37]	Análisis de datos de interacción de juegos cognitivos en aplicaciones de Smart TV para pacientes con Parkinson, Alzheimer y otros tipos de demencia	España	Smart TV	APP
[49]	Evaluación de la TVDi como herramienta de estimulación cognitiva en pacientes con Alzheimer en fase leve.	España	Smart TV	APP
Asistencia				
[36]	Tecnologías de asistencia para el Alzheimer y enfermedades relacionadas.	Canadá	Smart TV	FACT
[57]	Contribución a los sistemas de atención a personas mayores en el hogar a través de TVDi.	España	Smart TV	APP

3 Resultados

3.1 Clasificación por tipo de aplicación

En la figura 1 explícitamente se evidencia cómo las TIC incluida la TVDi postula el análisis de cómo mejorar una condición de deficiencia DTA en adultos mayores. Un 34,88% de tipo de experiencias categorizadas como análisis, donde se

promueven la inclusión social y el bienestar de los adultos mayores con el estudio de actividades y entornos que contribuyan a la participación mejorando así su calidad de vida y su sensación de pertenencia. Seguido se encuentra un 13,95% de publicaciones categorizadas como asistencia, donde se han explorado soluciones para ayudar a los adultos mayores con sus actividades diarias. Estas soluciones podrían incluir tecnologías y servicios que faciliten la realización de tareas cotidianas, como la administración de medicamentos, la movilidad y la gestión del hogar, mejorando así su independencia y autonomía. Las aplicaciones categorizadas como monitoreo (11,63%) aportan soluciones diseñadas para supervisar a los adultos mayores en sus actividades diarias. Estas soluciones podrían abarcar desde sistemas de alerta en caso de emergencias médicas hasta dispositivos de seguimiento que proporcionen tranquilidad a los familiares y cuidadores al mantenerlos informados sobre la ubicación y el estado del adulto mayor. Un 25,58% de las aplicaciones fueron categorizadas como de asistencia y monitoreo a la vez. Las aplicaciones categorizadas como lúdicas (11,63%) exploran la utilización de juegos como una forma de ayudar a los adultos mayores. Estos juegos están diseñados específicamente para abordar las necesidades y desafíos que enfrentan como deterioro cognitivo. Los juegos pueden mostrar los beneficios cognitivos y emocionales, manteniendo activas sus mentes y fomentando la interacción social. Finalmente, se halla una aplicación categorizada como compañía (2,33%), donde la investigación se basaba en otorgar un compañero virtual al adulto con pérdida de memoria y depresión.

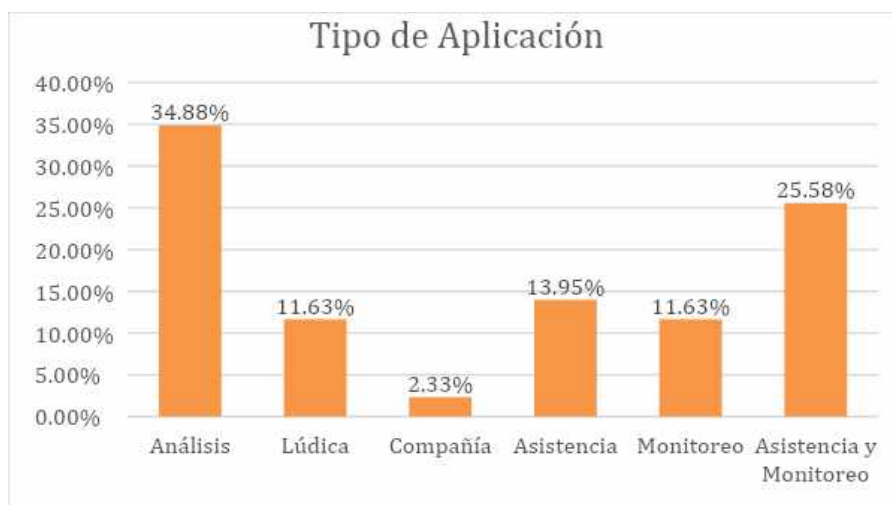


Fig. 1 Artículos por tipo de aplicación

3.2 Año de Publicación

La figura 2 muestra las publicaciones de artículos por año con tendencia a crecimiento del interés en el tema de tecnología y Alzheimer en los últimos años.

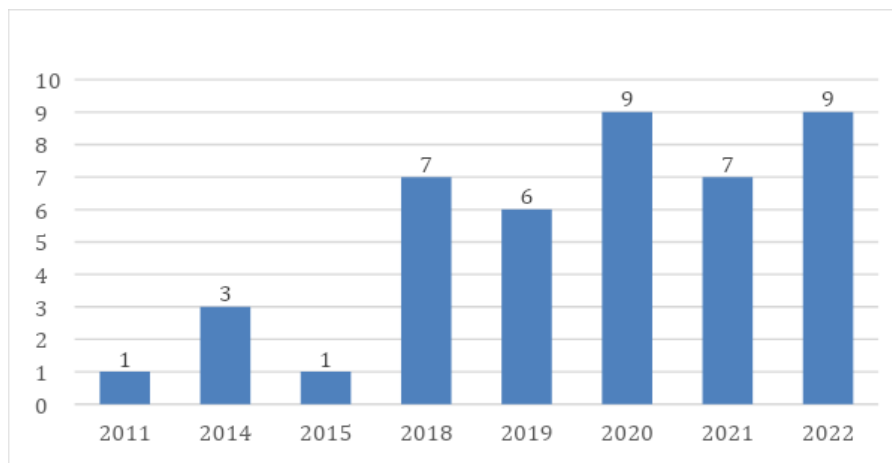


Fig. 2. Cantidad de estudios agrupados por año de publicación

3.3 Las TIC más aplicadas

La figura 3 muestra cómo las TIC han sido aplicadas en propuestas para mejorar la condición del Alzheimer en adultos mayores. En primer lugar, destaca el uso de computadoras como la herramienta principal, representando el 44,19% del total. A continuación, se encuentra el uso de dispositivos móviles y tablets, que alcanzan el 34,88%. Se identifica un 18,60% de propuestas que han empleado un Smart TV con contenidos interactivos. Uno de los artículos (2,33%) presenta el uso de un dispositivo específicamente diseñado como solución de interacción.

Dentro del conjunto de herramientas empleadas en el estudio, se destaca además la utilización de sensores en un 20,93% de las soluciones analizadas. Adicionalmente, se observa un interés del 2,78% en la aplicación de tecnologías de Realidad Virtual (RV) y Realidad Aumentada (RA). Los sensores se han empleado principalmente en el desarrollo de sistemas de monitoreo destinados a la captura de los movimientos de pacientes. En relación con la RV y RA, estas tecnologías se han explorado en la formulación de directrices para la creación de juegos dirigidos a personas con enfermedad de Alzheimer. Estos juegos se enfocan en la promoción de la interacción natural y la mejora de las habilidades cognitivas de los pacientes.

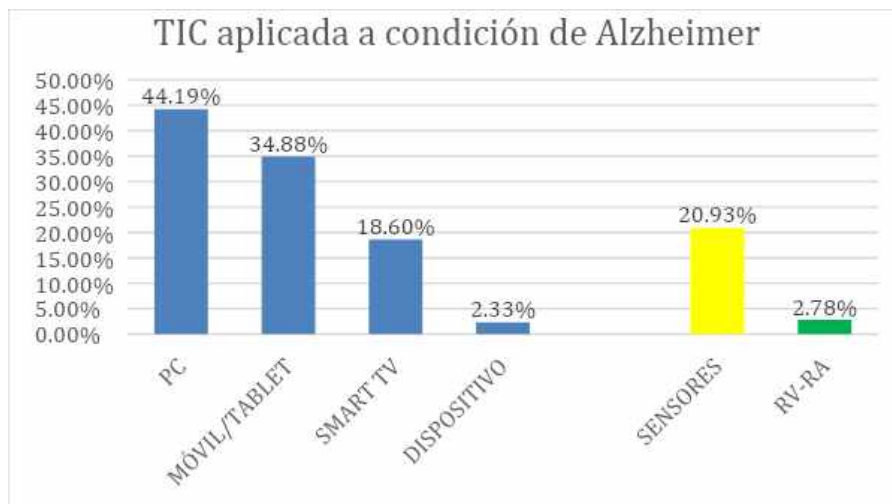


Fig. 3. Las TIC más utilizados en los estudios

3.4 Continentes donde se han desarrollado las propuestas de investigación

La figura 4 presenta los continentes donde tienen interés en el desarrollo de soluciones para la condición de Alzheimer. Siendo América y Europa donde se encuentran los estudios mayoritariamente. Países como Estados Unidos, Canadá, México, Brasil, Alemania, España, Suiza destacan en los artículos de este tema.

Los investigadores buscaban aportar con su trabajo investigativo a solucionar un tema latente como lo es la pérdida de memoria, con el fin de asistir tanto a quien lo cuida como al paciente.

Cabe destacar que, en las publicaciones revisadas, la mayoría (30) no presenta el país del grupo poblacional.

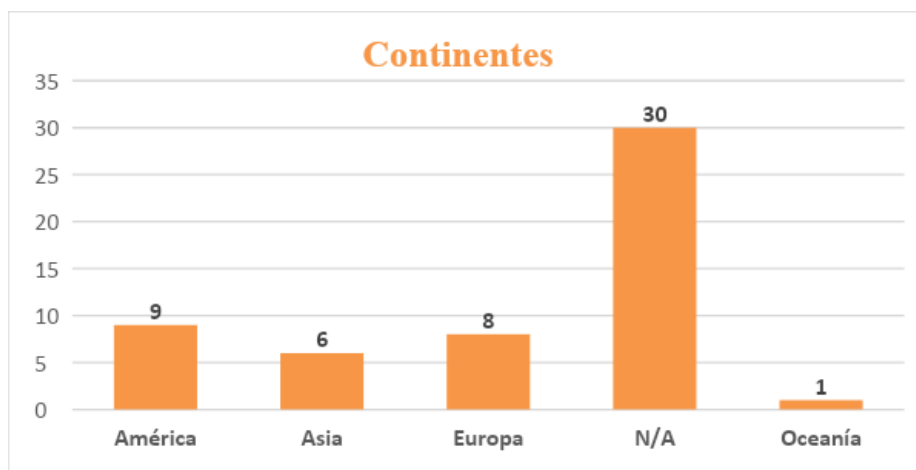


Fig. 4 Distribución de Continentes

3.5 Clasificación de estudios según el estado de sus desarrollos

En cuanto a la investigación se tiene que la mayoría de las propuestas, un 46,67%, han alcanzado la realización de prototipos, donde les falta la adición de elementos que lo hagan completamente funcional para comprobar su beneficio. El 44,44% hace referencia al uso de hardware y software pertinentes para que los adultos mayores acepten una solución tecnológica como ayuda a su problema de memoria y si están dispuestos o no a su uso prolongado con miras a mejorar su condición. Finalmente se tiene que el 8.89% ha logrado una implementación final, realizando pruebas, evaluaciones para retroalimentación a mejoras (figura 5).

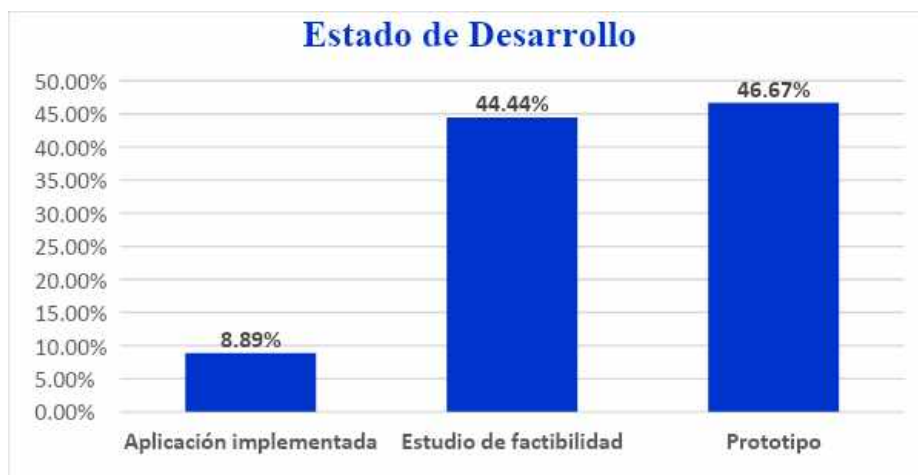


Fig. 5 Distribución por estado de desarrollo

3.6 Hallazgos más significativos sobre el uso de la TVDi

A continuación, se mencionan las aplicaciones para TVDi que fueron categorizadas como análisis. En el caso específico de personas con Alzheimer, la TVDi se ha mostrado como una herramienta viable y efectiva para la estimulación cognitiva, siendo bien aceptada tanto por los pacientes como por sus cuidadores [58]. Específicamente se encontró que la TVDi mejora la función cognitiva y disminuye la carga de cuidado para los cuidadores de pacientes con Alzheimer en fase leve [21]. Además, se observó una relación positiva entre la cantidad de tiempo de uso de la televisión interactiva y la mejora en la función cognitiva de los pacientes [41]. Un ejemplo del contexto mexicano [23] ha sido una plataforma electrónica basada en TVDi diseñada para ayudar a los adultos mayores a tener un mejor control de sus actividades diarias, como la toma de medicamentos y las citas médicas, lo que contribuye a una mejor calidad de vida al asegurar que no olviden tomar sus medicamentos ni faltar a sus citas médicas. Esta información se presenta a través de la televisión que tienen en sus casas, lo que crea un entorno amigable para los mayores. Asimismo, se ha desarrollado un episodio piloto de televisión de estimulación cognitiva [61] especialmente adaptado para satisfacer las necesidades

específicas de personas mayores con demencia, lo que representa un enfoque innovador en la atención y estimulación de este grupo de población.

En relación a las aplicaciones categorizadas como lúdicas se menciona que la TVDi como tecnologías de asistencia ha demostrado su efectividad para retrasar el deterioro cognitivo y mejorar la calidad de vida de los pacientes. En [37] se revela que los juegos cognitivos tienen un papel beneficioso para los pacientes, mejorando sus capacidades cognitivas y su socialización, además de contribuir a la detección temprana de la demencia y la toma de acciones adecuadas. Los profesionales resaltaron que un cribado automático [49], como los juegos cognitivos, resulta atractivo para los usuarios y puede ser una herramienta valiosa para la prescripción de tratamientos que retrasen la progresión de la enfermedad y aumenten la calidad de vida.

En relación a las aplicaciones para TVDi que fueron categorizadas como asistencia que se pueda dar a un paciente con déficit de memoria [36] se menciona al sistema HEREIAM desarrollado en una plataforma basada en TV para ayudar a las personas mayores a participar en la rutina diaria con tareas como comer, tomar pastillas. Asimismo en [57] se destaca el interés por probar el prototipo encaminado a saber de la salud del paciente con el “Abuelómetro”, que es información que pasa a los cuidadores o especialistas de salud donde pueden asistir al adulto mayor en caso de requerirse.

4 Conclusiones y trabajos futuros

La investigación presentada proporciona una revisión sistemática sobre cómo las TIC se aplican para enriquecer el apoyo brindado a pacientes que luchan contra el Alzheimer. Los resultados obtenidos de la revisión bibliográfica ilustran un abanico diverso de enfoques en el campo de la asistencia y el monitoreo, cuyo propósito radica en elevar la calidad de vida de estos individuos. Estas soluciones tecnológicas están diseñadas para aliviar a los adultos mayores en áreas como la memoria y las actividades cotidianas, como recordar la toma de medicamentos o mantenerse adecuadamente hidratados.

La tecnología adquiere un rol aún más influyente en la mejora de la vida de los adultos mayores afectados por el Alzheimer, acompañándolos en todas las etapas, desde las más tempranas hasta las más avanzadas de la enfermedad, de esta forma queda en evidencia el interés de investigaciones que destacan aplicaciones lúdicas, de compañía o como un complemento al trabajar propuesta de análisis y monitoreo a la vez para dar una asistencia integral.

Los desarrollos mayoritariamente hacen uso de PC, seguido del uso de dispositivos móviles. Se encontraron algunas propuestas que incorporan el uso de sensores para rastrear el comportamiento y la ubicación de los pacientes, mejorando así su seguridad y permitiendo una rápida respuesta en caso de desorientación. El uso de sensores a la hora de monitoreo es de vital importancia para la obtención de resultados según las lecturas presentadas. La Realidad Virtual y la Realidad Aumentada son tecnologías que comienzan a utilizarse en algunas de las propuestas analizadas.

Se encontraron algunos estudios sobre TVDi aplicada a este campo, los cuales han arrojado resultados prometedores y sugieren que esta puede ser una herramienta útil para retrasar el avance de la enfermedad y mejorar la calidad de vida de los pacientes.

Referencias

- [1] Abásolo, M. J., Rosado, M., Silva, T., Pina, J., Socorro, R., Kulesza, R., ... & Pesado, P. La Televisión Digital Interactiva para el mejoramiento de los pueblos latinoamericanos (2020).
- [2] Ahmed, Q. A., & Al-Neami, A. Q. (2020, July). A smart biomedical assisted system for alzheimer patients. In IOP Conference Series: Materials Science and Engineering (Vol. 881, No. 1, p. 012110). IOP Publishing.
- [3] Alberdi, A.; Weakley, A.; Schmitter-Edgecombe, M.; Cook, D. J.; Aztiria, A.; Basarab, A.; & Barrenechea, M.: Smart home-based prediction of multidomain symptoms related to Alzheimer's disease. *IEEE journal of biomedical and health informatics*, 22(6), pp.1720-1731 (2018).
- [4] Astell, A. J.; Panou, M.; Toulou, K.; Karavidopoulou, Y.; Cabrera-Umpiérrez, M. F.; Aldaz, E.; ... & Tabak, M.: Developing a pragmatic evaluation of ICTs for older adults with cognitive impairment at scale: the IN LIFE experience. *Universal access in the information society*, 21(1), 1-19 (2022).
- [5] Barreto, R.; Oliveira, F.; Sousa, A.; Cardoso and C. Duarte: Environment-aware system for Alzheimer's patients. *4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, 2014, pp. 300-303, doi: 10.1109/MOBIHEALTH.2014.7015970. (2014)
- [6] Birajdar, A. K., & More, P. (2018, August). Healthcare and security system for elderly and disabled people using ARM Microcontroller. In *2018 International Conference on Information, Communication, Engineering and Technology (ICICET)* (pp. 1-5). IEEE.
- [7] Boppana, L., Jain, V., & Kishore, R. (2019, July). Smart stick for elderly. In *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 259-266). IEEE.
- [8] Bouchard, B.; Bouchard, K.; & Bouzouane, A.: A smart cooking device for assisting cognitively impaired users. *Journal of reliable intelligent environments*, 6(2), 107-125 (2020)
- [9] Bustos Martínez, L. (2019). Mando a distancia y soledad: vejez y consumo televisivo en hogares unipersonales en el distrito de Usera.

- [10] Cáliz, D.; Ravivanpong, P.; Schankin, A.; Jadán-Guerrero, J.; Caraguay, W.; & Arellano, L.: Examining the usability of touchscreen gestures for adults with DS. *Journal of Reliable Intelligent Environments*, 7(4), 355-380 (2021).
- [11] Carillon, B., Dirian-Angeli, B., Breda, G., Benveniste, S., & Deparis, F. (2019). Design of new generation of serious games for patients with moderate cognitive impairment. In *Multi Conference on Computer Science and Information Systems, MCCSIS 2019-Proceedings of the International Conference on e-Health 2019* (pp. 245-249).
- [12] Casiddu, N.; Porfirione, C.; Monteriù, A.; & Cavallo, F. (Eds.). Ambient Assisted; Cleland, I., Zhang; S., Donnelly; M. P., ... & Tschanz, J.: Modelling mobile-based technology adoption among people with dementia. *Personal and Ubiquitous Computing*, 26(2), 365-384 (2022)
- [13] Chaurasia, P., McClean, S., Nugent, C. D., Cleland, I., Zhang, S., Donnelly, M. P., ... & Tschanz, J. (2022). Modelling mobile-based technology adoption among people with dementia. *Personal and Ubiquitous Computing*, 26(2), 365-384.
- [14] Cortés, R. P. B.; Espinoza, R. L.; Navarrete, P. A.; González, Y. C.; & Covarrubias, V. J. T. Meyotl: A Web-App Prototype for the Geolocation of People with Cognitive Impairment. In 2019 *International Conference on Inclusive Technologies and Education (CONTIE)* (pp. 85-855). IEEE. (2019, October)
- [15] Crocetta, T. B.; de Araújo, L. V.; Guarnieri, R.; Massetti, T.; Ferreira, F. H. I. B.; De Abreu, L. C.; & de Mello Monteiro, C. B.: Virtual reality software package for implementing motor learning and rehabilitation experiments. *Virtual Reality*, 22(3), 199-209 (2018)
- [16] D. Machesney; S. S. Wexler; T. Chen and J. F. Coppola.: "Gerontechnology Companion: Virtual pets for dementia patients," *IEEE Long Island Systems, Applications and Technology (LISAT) Conference*, pp. 1-3, doi: 10.1109/LISAT.2014.6845226 (2014).
- [17] Dinuwanthi, I.; Thilakarathna, H.; Sripadi, V.; Herath, D.; & Ragel, R.: Revealing MicroRNA Biomarkers for Alzheimer's Disease Using Next Generation Sequencing Data. In 2021 *10th International Conference on Information and Automation for Sustainability (ICIAfS)* (pp. 117-122). IEEE (2021, August)
- [18] Ferm, U.; Ekström, A.; Larsson, E.; & Samuelsson, C.: Tablet computer-supported conversation between people with dementia and their carers: technology as interactional focus. *Universal Access in the Information Society*, 20(4), 753-765 (2021)
- [19] Francillette, Y.; Boucher, E.; Bier, N.; Lussier, M.; Bouchard, K.; Belchior, P.; & Gaboury, S.: Modeling the behavior of persons with mild cognitive impairment or Alzheimer's for intelligent environment simulation. *User Modeling and User-Adapted Interaction*, 30(5), 895-947 (2020)
- [20] Fredericks, E. M.; Bowers, K. M.; Price, K. A.; & Hariri, R. H. Cal: A smart home environment for monitoring cognitive decline. In 2018 *IEEE 38th*

- International Conference on Distributed Computing Systems (ICDCS)* (pp. 1500-1506). IEEE (2018, July)
- [21] García-Cabot, A., Garrido, M., & Perea, M. V. Diseño de un sistema interactivo de televisión para el entrenamiento cognitivo de pacientes con Alzheimer. *Revista de Investigación en Tecnología e Informática*, 1(1), 25-33. (2018).
- [22] Gimenez Manuel, J. G.; Augusto, J. C.; & Stewart, J. AnAbEL: Towards empowering people living with dementia in ambient assisted living. *Universal Access in the Information Society*, 1-20 (2020)
- [23] González-Gómez, A. A., & Jiménez-Comrie, K. (2006). La televisión digital interactiva y sus aplicaciones educativas. *Comunicar: Revista Científica de Comunicación y Educación*, 13(26), 93-101.
- [24] Harish, M. V.; Dinesh, C.; Sasikala, S.; & Kumar, A.: Alzheimer's Disease Prediction Using Machine Learning Methodologies. In 2022 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE (2022, January)
- [25] Hoe, Z. Y.; Lee, I. J.; Chen, C. H.; & Chang, K. P.: Using an augmented reality-based training system to promote spatial visualization ability for the elderly. *Universal Access in the Information Society*, 18(2), 327-342 (2019)
- [26] Høeg, E. R.; Bruun-Pedersen, J. R.; Cheary, S.; Andersen, L. K.; Paisa, R.; Serafin, S.; & Lange, B.: Buddy biking: a user study on social collaboration in a virtual reality exergame for rehabilitation. *Virtual Reality*, 1-18 (2021)
- [27] Johansson, S.; Gulliksen, J.; & Gustavsson, C.: Disability digital divide: the use of the internet, smartphones, computers and tablets among people with disabilities in Sweden. *Universal Access in the Information Society*, 20(1), 105-120 (2021)
- [28] Kanno, K. M.; Lamounier, E. A.; Cardoso, A.; Lopes, E. J.; & de Lima, G. F. M.: Augmented reality system for aiding mild alzheimer patients and caregivers. In 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 593-594). IEEE. (2018, March)
- [29] Karppinen, M. S. L.; Karppinen, J.; & Halonen, R.: A Concept to Improve Care for People with Dementia. In *Advances in Computer Communication and Computational Sciences* (pp. 603-618). Springer, Singapore. (2019)
- [30] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- [31] Ladly, M. J; K., Farrelly, G.; Micak, K.; Penn, G.; & Rudzicz, F.: Reality recalled: Elders, memory and VR. In 2017 23rd International Conference on Virtual System & Multimedia (VSMM) (pp. 1-9). IEEE. (2017)
- [32] Ladly, M.; & Chadha, K.: Postcard memories: a virtual/tangible memory sharing application for adults with early-stage dementia (ESD). *Virtual Creativity*, 10(1), 79-88 (2020)

- [33] Li, C.; Hu, J.; Hengeveld, B.; & Hummels, C.: Supporting intergenerational memento storytelling for older adults through a tangible display: a case study. *Personal and Ubiquitous Computing*, 1-25 (2020)
- [34] Lin, Y. W.; & Lin, C. Y.: An interactive real-time locating system based on bluetooth low-energy beacon network. *Sensors*, 18(5), 1637 (2018)
- [35] Liu, M. (2019, December). The Research of Gait Recognition Based on Deep Learning: A Case Study of the Missing Elderly. In IOP Conference Series: Materials Science and Engineering (Vol. 677, No. 3, p. 032072). IOP Publishing.
- [36] López-de-Ipiña, K., Alonso-Revuelta, B., Barroso, N., & Solé-Casals, J. Ambient assisted living technologies for Alzheimer's disease and related disorders. *Journal of Alzheimer's Disease*, 52(4), 1307-1316. (2016).
- [37] López, J. P., Moreno, F., Popa, M., Hernández-Peñaloza, G., & Álvarez, F. (2019). Data analysis from cognitive games interaction in Smart TV applications for patients with Parkinson's, Alzheimer's, and other types of dementia. *AI EDAM*, 33(4), 442-457.
- [38] Machesney, D., Wexler, S. S., Chen, T., & Coppola, J. F. (2014, May). Gerontechnology companion: Virtual pets for dementia patients. In IEEE long island systems, applications and technology (LISAT) conference 2014 (pp. 1-3). IEEE.
- [39] M. J. Manu; P. S. Akshay; G. Kattukaran; Glenchrist, A. Anto and V. P. Madhanmohan.: "A Smart Wearable Device for Tracking and Fall Detection for Geriatric People, Alzheimer Patients and Children with Autism," 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), pp. 333-337, doi: 10.1109/ICEECCOT52851.2021.9708064 (2021)
- [40] Martínez-Sánchez, F., Meilán, J. J. G., Pérez, E., Franco-Martín, M. A., & Carro, J. Interactive TV for people with Alzheimer's disease: An enriched entertainment device for cognitive stimulation. *PsychNology Journal*, 14(2), 125-139. (2016).
- [41] Martínez-Sánchez, F., Meilán, J. J. G., Pérez, E., Franco-Martín, M. A., & Carro, J. Interactive television as a tool for cognitive stimulation in people with Alzheimer's disease: A feasibility study. *Journal of Alzheimer's Disease*, 55(3), 1285-1294. (2017).
- [42] Matsangidou, M.; Frangouides, F.; Schiza, E.; Neokleous, K. C.; Papayianni, E.; Xenari, K.; ... & Pattichis, C. S.: Participatory design and evaluation of virtual reality physical rehabilitation for people living with dementia. *Virtual Reality*, 1-18 (2022)
- [43] Megalingam, R. K.; Pocklassery, G.; Jayakrishnan, V.; Mourya, G.; & Thulasi, A. A.: Smartphone based continuous monitoring system for home-bound elders and patients. In 2014 *International Conference on Communication and Signal Processing* (pp. 1173-1177). IEEE. (2014, April)
- [44] MohanaPriya, D.; Deepika, V.; Priya, M. S.; & Yogeswari, C. S.: A Real Time Support System to Impart Medicine Using Smart Dispenser. In 2020

- International Conference on System, Computation, Automation and Networking (ICSCAN)* (pp. 1-10). IEEE. (2020, July)
- [45] Monteagudo, J. L. (Capacidades y oportunidades de innovación en TIC para Alzheimer, (2013).
- [46] Muñoz-Montoro, A. J., Cano-de-la-Cuerda, R., & Alguacil-Diego, I. M. Interactive television system as a tool for cognitive estimation in elderly people: A systematic review. *Journal of Medical Systems*, 42(9), 173. (2018).
- [47] Ng, K.-D.; Mehdizadeh, S.; Iaboni, A.; Mansfield, A.; Flint, A.; & Taati, B.: Measuring Gait Variables Using Computer Vision to Assess Mobility and Fall Risk in Older Adults with Dementia. *IEEE Journal of Translational Engineering in Health and Medicine*, 1–1 (2020)
- [48] Obaid, A. K., Abdel-Qader, I., & Mickus, M. (2018, November). Automatic Food-Intake Monitoring System for Persons Living with Alzheimer's-Vision-Based Embedded System. In 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 580-584). IEEE.
- [49] Oña, E. D., Camacho, J. E., Perales, F. J., Verdú, E., Valls-Solé, J., Reñé, R. M., Navarro, M. D., & Alcañiz, M. Interactive Television as a Tool for Cognitive Stimulation of Alzheimer's Disease Patients: A Pilot Study. *Journal of Alzheimer's Disease*, 36(4), 883-888 (2013).
- [50] Organización Mundial de la Salud. (15 de marzo de 2023). Demencia. Recuperado de: <https://www.who.int/es/news-room/fact-sheets/detail/dementia>
- [51] Ozdemir, D.; Cibulka, J.; Stepankova, O.; & Holmerova, I.: Design and implementation framework of social assistive robotics for people with dementia-a scoping review. *Health and Technology*, 11(2), 367-378 (2021)
- [52] Palacio, R. R.; Acosta, C. O.; Cortez, J.; & Morán, A. L.: Usability perception of different video game devices in elderly users. *Universal Access in the Information Society*, 16(1), 103-113 (2017)
- [53] Pirhonen, J.; Melkas, H.; Laitinen, A.; & Pekkarinen, S.: Could robots strengthen the sense of autonomy of older people residing in assisted living facilities? - A future-oriented study. *Ethics and Information Technology*, 22(2), 151-162 (2020)
- [54] Qiu, S.; An, P.; Kang, K.; Hu, J.; Han, T.; & Rauterberg, M.: Investigating socially assistive systems from system design and evaluation: A systematic review. *Universal Access in the Information Society*, 1-25 (2021).
- [55] Renyi, M., Gaugisch, P., Hunck, A., Strunck, S., Kunze, C., & Teuteberg, F. (2022). Uncovering the Complexity of Care Networks—Towards a Taxonomy of Collaboration Complexity in Homecare. *Computer Supported Cooperative Work (CSCW)*, 1-38.
- [56] Ruttimann, C., Karatli, E., Norões, A. D., Rebello, L., & Amado, G. (2019, July). Sequaca—Cognitive Rehabilitation Game Kit for Treatment of Neurocognitive Disorders. In *International Conference on Applied Human Factors and Ergonomics* (pp. 911-920). Springer, Cham.

- [57] Santana Mancilla, P. C. (2022). Contribución a los sistemas de atención a personas mayores en el hogar a través de TVDi e IdC (Doctoral dissertation, Enxeñaría telemática).
- [58] Santana-Mancilla, P. C., Contreras-Castillo, J., & Anido-Rifón, L. E. (2019, September). Designing for social TVDi: improving the shared experience of home care systems. In Proceedings of the IX Latin American Conference on Human Computer Interaction (pp. 1-4).
- [59] Soufian, M., Nefti-Mezian, S., & Drake, J. (2020). Toward Kinecting cognition by behavior recognition-based deep learning and big data. *Universal Access in the Information Society*, 1-19.
- [60] Stavropoulos, T. G., Lazarou, I., Strantsalis, D., Nikolopoulos, S., Kompatsiaris, I., Koumanakos, G., ... & Tsolaki, M. (2020, September). Human factors and requirements of people with mild cognitive impairment, their caregivers and healthcare professionals for eHealth systems with wearable trackers. In 2020 IEEE International Conference on Human-Machine Systems (ICHMS) (pp. 1-6). IEEE.
- [61] Streater, A., Yates, L., Orrell, M., Rosen, J., Taylor Smith, A., & Schneider, J. (2020). Interacting with television in one's own home: The development of a cognitive stimulation television pilot episode for older people with dementia (Innovative Practice). *Dementia*, 19(8), 2881-2888.
- [62] Wexler, S. S., Drury, L. J., Coppola, J. F., Tschinkel, B. J., & Thomas, B. A. (2011, May). Service-learning computing courses assist with technology needs in community-based organizations serving older adults. In 2011 IEEE Long Island Systems, Applications and Technology Conference (pp. 1-6). IEEE.

Técnicas de Visión Artificial utilizadas en la detección y seguimiento de una pelota en imágenes de un partido de fútbol para accesibilidad de personas con discapacidad visual a la TV.

Diego Villamarín^{1,2}[0000-0001-9175-9570], Lothar Tierra²[0000-0002-3215-4704], David Moya²[0000-0002-0361-3799], José Manuel Menéndez¹[0000-0003-0584-2250], Julio Larco²[0000-0003-4789-5983]

¹ Universidad Politécnica de Madrid, Madrid, España

df.villamarin@alumnos.upm.es, jmm@gatv.ssr.upm.es

² Universidad de las Fuerzas Armadas - ESPE, Sangolquí, Ecuador

{dfvillamarin, latierra, dsmoyal, jclarco}@espe.edu.ec

Abstract. Las tareas de seguimiento y detección de una pelota en una cancha de fútbol son de gran importancia para que las personas con discapacidad visual se puedan incluir de manera autónoma en la transmisión de este famoso deporte mediante la TV, sin embargo, es un completo desafío debido a muchos inconvenientes como la sombra e iluminación de la pelota, tamaño, superposición de la pelota con otros objetos como jugadores o líneas de campo, entre otros. Aunque existen trabajos relacionados que resuelven parcialmente la problemática presentada, estos no contienen información completa o presentan la información de manera dispersa. A razón de este problema, este artículo tiene como principal objetivo proporcionar una recopilación de las técnicas, herramientas y algoritmos de la visión artificial disponibles actualmente, que nos permita adelante implementar una solución de detección y seguimiento de una pelota en imágenes de un partido de fútbol con baja latencia y alta confiabilidad. Además, se presenta una explicación de las ventajas y desventajas de las diferentes técnicas encontradas para realizar las tareas de detección y seguimiento de objetos, y una comparación entre ellas. Finalmente, se presenta un camino o una ruta prometedora para la investigación y el desarrollo de trabajos futuros hacia el campo de la visión artificial con objetivo de ayudar a personas con discapacidad visual, y que los investigadores se puedan comprometer a brindar soluciones a desafíos no resueltos de este tema.

Keywords: Técnicas, Detección de Objetos, Visión Artificial, Personas con discapacidad visual.

1 Introducción

La discapacidad visual presenta un reto para el desarrollo de las funciones y actividades del ser humano, y una de las que más prevalece a nivel mundial. En el trabajo de

Hurtado y Agudelo [1] se menciona que, uno de los objetivos principales de cada nación es lograr una inclusión exitosa de las personas con discapacidad visual, una participación de derechos civiles y de igualdad de oportunidades. Actualmente, la medicina aún no cuenta con un método efectivo para la restauración total o parcial de la vista por lo que, presenta barreras que limitan la inclusión de las personas con discapacidad visual, por lo que en la mayoría de veces, estas personas cumplen un rol pasivo en la sociedad; a esto se suma la falta de tecnologías asistidas y falta de información técnica que permitan facilitar la inclusión de dichas personas, según González et al. [2]. La participación o inclusión en la sociedad, representa un cambio positivo para la persona con discapacidad, porque le permite acceder a una mayor interacción con el entorno y con posibilidades de participar en eventos sociales y laborales de manera autónoma.

Existe una amplia gama de aplicaciones en las que se puede utilizar visión artificial, que sirven de ayuda para la inclusión de las personas con discapacidad visual para la identificación y seguimiento de objetos, como el análisis de vídeos deportivos donde se requiere saber la ubicación exacta de los jugadores y de la pelota, según Najeeb y Ghani [3].

Sin embargo, la falta de conocimiento de la existencia de técnicas, herramientas y algoritmos disponibles basados en visión artificial para detección y seguimiento de objetos, es la principal limitante para la evolución y desarrollo de nuevas tecnologías que permitan su implementación en dispositivos hápticos que sirvan de ayuda a la inclusión de las personas no videntes. En el presente trabajo se muestra el estado del arte los diferentes métodos existentes y adaptados para detección y seguimiento de un balón en un partido de fútbol o similares, mediante el procesamiento digital de imágenes secuenciales y vídeos; con la finalidad de brindar una solución factible a la problemática que aún no se logró solucionar en un trabajo anterior [4], en el que se presentó un guante háptico que permite la accesibilidad a la TV para personas con deficiencias visuales que no han sido incluidas en desarrollo de TV inmersiva y TV interactiva.

2 Trabajos relacionados

En este documento, nos enfocamos en realizar una revisión de las técnicas y métodos de visión artificial que pueden ser utilizadas en el seguimiento y detección de una pelota en un vídeo de un partido de fútbol de manera confiable y fiable. Esta sección es muy importante porque es la base de nuestra investigación, se busca información relevante de otros trabajos investigativos que han sido parte de interés por parte de la comunidad científica.

En el trabajo de Yuan et al. [5], se presenta un análisis respecto al progreso y los desarrollos recientes en las técnicas de análisis visual útiles para Machine Learning. Además, clasifica a las técnicas en tres principales grupos según la etapa de análisis correspondiente: técnicas antes, durante y después de la construcción de los modelos de aprendizaje automático; en cada una de estas etapas, se busca mejorar la calidad de

los datos completando los atributos faltantes y corregir las etiquetas de datos incorrectas; se menciona que el futuro de las técnicas empleadas para la visión artificial consta de redes neuronales que se caracteriza por la selección y construcción de las características de los objetos de interés.

En el trabajo de Macaulay y Shafiee [6], se presenta una categorización distinta de los diferentes algoritmos de aprendizaje automático: algoritmos de aprendizaje supervisado, no supervisado, aprendizaje por refuerzo y aprendizaje profundo (Deep Learning - DL).

El aprendizaje supervisado se basa en predicciones que a su vez se pueden categorizar en clasificación y regresión de los datos obtenidos, para ello se separa los datos en un conjunto de entrenamiento y en otro conjunto de prueba. El aprendizaje no supervisado es aquel que realiza una búsqueda de rasgos característicos similares que puedan usarse para agrupar los datos por categorías o clases. El aprendizaje por refuerzo es una combinación y tiene características de los dos aprendizajes antes mencionados, debido que no requiere que el sistema aprenda de los datos, sino que es el resultado de una retroalimentación de los datos que procesa. Y finalmente se tiene el aprendizaje profundo que requiere de una gran cantidad de datos para que los sistemas o algoritmos basados en este tipo de aprendizaje puedan producir valores, predicciones o predicciones correctas. En la **Tabla 1**, se presenta un resumen los algoritmos que se basan en los aprendizajes previamente descritos en el trabajo [6].

Tabla 1: Categorización de los algoritmos dependiendo el tipo de aprendizaje

	Tipo de Aprendizaje				
	Supervisado		No Supervisado	Por Refuerzo	Profundo
	Basado en la regresión	Basado en la clasificación			
Algoritmos	<ul style="list-style-type: none"> • Regresión lineal • Bosque aleatorio 	<ul style="list-style-type: none"> • Regresión logística • Máquina de vectores de soporte (VSM) • Árbol de decisión • Red neuronal artificial (ANN) • K-nearest neighbors (KNN) 	<ul style="list-style-type: none"> • Agrupamiento de K-means • Agrupamiento jerárquico • Detección de anomalías • Análisis de componentes principales (PCA) • Descomposición de valores singulares • Algoritmo a priori 	<ul style="list-style-type: none"> • Algoritmo SARSA-Lambda • Deep Q Network (DQN) • Deep Deterministic Policy Gradient (DDPG) • Asynchronous Advantage Actor-Critic (A3C) 	<ul style="list-style-type: none"> • Red neural profunda (DNN) • Redes de creencia profunda (DBNs) • Redes neuronales convolucionales recurrentes (RCNN) • Fast RCNN • Faster RCNN • Mask RCNN • Detector de disparo único (SSD) • You Only Look Once (YOLO)

En el trabajo de Tan et al. [7], se realiza una revisión bibliográfica exhaustiva y se tiene que los algoritmos de aprendizaje profundo de detección y seguimiento de objetos en imágenes se pueden dividir en dos grupos basándose en la cantidad de veces que la

imagen es procesada para obtener valores, predicciones y detecciones, los que son de una etapa, y los que son de dos etapas. El trabajo de Tan et al., se centra en el estudio de los algoritmos de una sola etapa, donde se menciona principalmente RetinaNet, SSD o YOLO que se caracterizan por utilizar una única red para predecir de manera directa los cuadros delimitadores de los objetos, aunque este tipo de algoritmos presenta una gran ventaja respecto al poco tiempo de procesamiento, lo que es ideal para una aplicación en tiempo real, la precisión del algoritmo para la detección de objetos pequeños no es la adecuada. Además, se realizó una comparación entre los modelos antes descritos, como se observa en **Tabla 2**.

Tabla 2: Comparación de modelos Deep Learning [7]

Evaluación de modelos Deep Learning					
Algoritmo	Precisión	Recuperación	Puntuación F1	MAP	FPS
RetinaNet	64,98%	83,86%	73,26%	82,89%	<30
SSD	63,69%	88,89%	74,21%	82,71%	32
YOLOv3	69,65%	80,67%	74,77%	80,69%	51

Para la evaluación se consideran 4 categorías o resultados de las pruebas de salida, en comparación a las etiquetas verdaderas, las cuales son: verdaderos positivos (TP), falsos positivos (FP), falsos negativos (FN) y verdaderos negativos (TN), donde los píxeles que se encuentren dentro de la Región de Interés (ROI) se consideran positivos, y los píxeles que se encuentren fuera de la región se consideran negativos, según Mendoza y Pérez [8]. Las métricas de evaluación incluyen recuperación, precisión, puntaje F1, precisión media promedio (MAP) y cuadros por segundo (FPS).

- **Precisión.** - es la relación de los objetos detectados correctamente y el número de todos los objetos detectados.

$$P = \frac{TP}{TP + FP}$$

- **Recuperación.** – es la relación entre el número de objetos detectados correctamente y todos los objetos en el conjunto de la muestra.

$$R = \frac{TP}{TP + FN}$$

- **Puntuación F1.**- es el promedio armónico ponderado de precisión y recuperación.

$$F1 = \frac{2PR}{P + R}$$

- **MAP.** - es numéricamente igual al valor de la suma AP (precisión en todos los elementos de una categoría) en todas las categorías, y este valor se utiliza como herramienta para evaluar el modelo de manera general.
- **FPS.** - es la cantidad de imágenes que se pueden procesar en un segundo, y se considera que a mayor o igual 30 FPS se tiene un sistema que trabaja en tiempo real.

3 Metodología

Para recopilar la información necesaria para este artículo, se aplicó la metodología de recuperación documental y la técnica de muestreo por cuotas de la información recuperada, debido que esta permite separar la información seleccionada por estratos o por determinadas características [9], lo que hace que la información obtenida pueda ser clasificada de acuerdo a nuestro interés. Se realizó una búsqueda de documentación bibliográfica en la base de datos Web Of Science, Springer e IEEE Xplore. Las fuentes de información se basaron en artículos científicos y en trabajos de titulación de grado con fechas de publicación hasta hace 5 años atrás, con el objetivo de obtener tecnologías vigentes, que están en constante actualización y sean tendencias futuras.

Se empezó con una búsqueda preliminar en las bases de datos antes mencionadas, donde se presentaron los siguientes resultados: (a) en la base de datos de Springer se utilizaron las palabras claves “artificial vision + survey + ball detection” (66 resultados), (b) en la base de datos de Web Of Science, se aplicaron las palabras claves a la búsqueda “object detection + artificial vision+ techniques” (143), y finalmente se realizó la búsqueda en la revista IEEE Xplore con las palabras claves “object detection + techniques+ artificial vision” (115 resultados). Se obtuvo un total de 314 documentos bibliográficos, de los cuales, debido al empleo de palabras claves similares, se discriminaron los documentos duplicados quedando así un total de 186 resultados, de los artículos obtenidos en total, se eliminaron aquellos que no correspondían a los objetivos de este documento, quedando un total de 105 resultados. El siguiente paso en la metodología fue revisar los resúmenes de los artículos candidatos para determinar si se referían a técnicas de análisis visual para la detección de objetos. Si el título y el resumen no proporcionaban información clara, se revisaba el texto completo para tomar una decisión final, además se incluyeron los artículos en este documento a aquellos que cumplían con los siguientes criterios:

- En su prioridad, describan técnicas dedicadas especialmente a la detección y seguimiento de objetos redondos, con la finalidad que, sean adecuados para la aplicación a una pelota de fútbol.
- Existan aplicaciones realizadas con las técnicas descritas, por la razón que, demuestran que han sido estudiadas, aplicadas y verificadas por la comunidad científica en el contexto de interés.

- Trabajos en los cuales se describan ideas o aplicaciones novedosas que aporten a la inclusión de las personas con discapacidad visual a realizar tareas de manera autónoma, además que estas ideas no hayan sido mencionadas en otros artículos.
- Técnicas que se puedan desarrollar mediante entornos de desarrollo (plataformas de programación) que permitan un desarrollo ágil del código que se encargará de procesar las imágenes y arrojar los resultados de interés de manera confiable y fiable. Además, que sean aplicables para un sistema tiempo real, para esto se consideró el valor de los fps (frame por segundo) que se refiere a la cantidad de imágenes procesadas por un segundo; un sistema que aplique una técnica con un valor de fps mayor a 30 segundos es considerado de tiempo real.

Aplicando estos criterios se obtuvo un total de 20 trabajos, de estos trabajos se realizó una revisión y lectura exhaustiva para elaborar un resumen con lo primordial de cada técnica, con el fin de tener una idea clara de sus funciones principales y características relevantes que sean de ayuda para el investigador al momento de escoger la técnica apropiada para el desarrollo de su aplicación en interés. Luego de analizar cada uno de los trabajos bibliográficos obtenidos, se llegó a la decisión de elaborar una taxonomía desde el punto de vista de la tecnología aplicada en cada uno de los métodos. Utilizando estos criterios, se presenta la clasificación de los métodos en 4 grupos: (a) métodos clásicos de procesamiento de imágenes, (b) métodos tradicionales basados en aprendizaje automático (Machine Learning), (c) métodos basados en aprendizaje profundo (Deep Learning) y (d) los potenciales métodos a utilizar.

Según Rashid M. et al. [10] dentro de la clasificación, seguimiento y detección de objetos una de las complicaciones críticas que puede presentarse es el hecho de no escoger un método apropiado de extracción de características. Una selección incorrecta sin duda afectará la precisión de la clasificación y seguimiento provocando errores en la toma de decisiones. Por lo tanto, es necesario realizar una intensa investigación sobre la eficiencia de varios métodos de seguimiento y detección de objetos. Un método de extracción de características puede funcionar bien con un algoritmo de seguimiento y detección de partículas, pero puede no funcionar bien si se utiliza en otro entorno o utilizando una herramienta diferente, es decir, es necesario realizar una comparativa de los resultados que se han obtenido en diferentes ámbitos de aplicación sean similares o diferentes para así realizar un estudio analítico detallado de cuál es la técnica adecuada para lo que se desea realizar.

3.1 Herramientas para el empleo y desarrollo de algoritmos orientados a visión artificial

Como punto de partida es de gran importancia conocer que, existen algunos entornos de desarrollo para el procesamiento de imágenes y detección de objetos en imágenes o videos empleando visión artificial, cada uno con sus propias características y ventajas, por lo que es importante evaluar y escoger la adecuada para la aplicación requerida, a continuación, se realiza una descripción de algunos de los más relevantes y que han dado buenos resultados para el empleo de visión artificial:

A. Matlab

Matlab, es un entorno de desarrollo con su propio lenguaje nativo y su propia IDE, cuenta con algunas herramientas para el análisis matemático en ingeniería. Es una herramienta apropiada para realizar algoritmos de visión artificial, debido a que puede procesar imágenes, comparar los colores, discriminar rasgos característicos como el fondo o morfología de un objeto y detección de este [11].

Adicionalmente MATLAB cuenta con una serie de Toolbox que proporcionan apps, algoritmos y redes entrenadas que puede utilizar para desarrollar prestaciones de visión artificial [12].

B. PyTorch

PyTorch es una biblioteca de aprendizaje automático y aprendizaje profundo de código abierto que proporciona un uso fluido de las GPU y una plataforma para el aprendizaje profundo que proporciona la máxima flexibilidad y velocidad [13].

C. TensorFlow

TensorFlow es una interfaz para el desarrollo de algoritmos de aprendizaje automático y una implementación para ejecutar dichos algoritmos, desde dispositivos móviles como teléfonos y tabletas hasta sistemas distribuidos a gran escala de cientos de máquinas y miles de dispositivos computacionales como tarjetas GPU [14].

D. OpenCV

OpenCV, es una biblioteca de código abierto para procesamiento de imágenes o videos mediante el uso de redes neuronales para visión artificial con implementaciones que abarcan más de 2500 algoritmos [11].

E. Lebel DB

Es una librería de almacenamiento rápido de pares clave-valor, esta base de datos se utiliza para almacenar las imágenes usadas para el entrenamiento de forma que a cada imagen se le asigna una etiqueta en función de si hay un coche o no [14].

F. Keras

Keras es una biblioteca de código abierto cuyo objetivo es acelerar la creación de redes neuronales: para ello, Keras no funciona como un framework independiente, sino como una interfaz de uso intuitivo (API) que permite acceder a varios frameworks de aprendizaje automático y desarrollarlos [15].

G. LabView

LabView, contiene un entorno de programación gráfico, posee la gran ventaja de contar con servidores orientados al procesamiento en tiempo real, además es un entorno licenciado y de alto costo computacional [11].

3.2 Técnicas para seguimiento de objetos mediante algoritmos orientados a visión artificial

La detección de objetos es una tarea larga y difícil en la visión por ordenador [16]. Como una de las tareas principales de la visión por ordenador, el objetivo final de la detección de objetos es dar las clases y ubicación de los objetos. Por lo tanto, la detección de objetos es la base de muchas otras tareas de visión artificial, como seguimiento de objetos [17] y la comprensión de escenas [18]. La detección de objetos y sus métodos evolucionan continuamente desde que fueron propuestos, a continuación, se busca resumir de manera exhaustiva el desarrollo de los métodos de detección de objetos de los últimos años y agrupar los métodos tradicionales basados en características manuales o basados en aprendizaje profundo [19].

A. YOLO

YOLO es una técnica o algoritmo que utiliza redes neuronales convolucionales para la detección de objetos en imágenes, e incluso puede ser en tiempo real. Esta arquitectura fue presentada por Joseph Redmon en [20], actualmente existen cinco versiones además de Fast o Tiny YOLO. La estrategia del modelo es la detección de objetos en una sola pasada por la red, originando así su nombre: You Only Look Once (Sólo se mira una vez), lo que lo hace más rápido que otros algoritmos, aunque no es el más preciso.

La diferencia de YOLO con redes neuronales convolucionales ordinarias se basa en que estas redes ordinarias no procesan todo el fotograma, sólo a regiones individuales y esto las hace más lentas. Como YOLO puede detectar varios objetos en una imagen, esto significa que divide la imagen en regiones y además de predecir la también predice las probabilidades de cada región, creando un mapa de probabilidades de clase [21]. Por último, se aplican filtros para eliminar los recuadros delimitadores con probabilidad inferior a un umbral y las que se solapan, obteniendo así las detecciones de los objetos de la imagen, estos modelos pueden predecir coordenadas de posición [22].

B. Backpropagation

En el trabajo de Rocha, C. et al. [23] dice que Backpropagation es una red neuronal de aprendizaje multicapa, que se puede entrenar y es capaz de aprender el algoritmo correcto relacionado con la tarea solicitada sin que alguien lo halla deducido por anticipado.

C. Kernelized Filter (KCF)

En el trabajo de Montoya [24], explica que los filtros de correlación de Kernelizados (KCF), es un seguidor de objetos en imágenes procesadas que utiliza propiedades de la matriz circulante para mejorar la velocidad de procesamiento; su principio básico es: al tener un parche en la imagen de interés, donde se tiene el objeto deseado, un clasificador es entrenado para que pueda discriminar la morfología del objeto de interés con el entorno.

D. Multiple Instance Learning (MIL)

En el trabajo de Montoya [24], explica que este algoritmo es una mejora del anterior y se basa en tres componentes esenciales: representación de la imagen, modelo de la apariencia y modelo de movimiento. La entrada para este algoritmo es un cuadro de video k el cual debe ser segmentado en regiones que componen dicha imagen. El modelo de apariencia para este algoritmo se compone por un discriminador, el cual permite el retorno de una validación de si el objeto se encuentra en el cuadro o región actual. Para cada Frame procesado, se corta un conjunto de pedazos de la imagen que está dentro de un radio de búsqueda de la ubicación actual del objeto de seguidor; posteriormente se actualiza el seguidor (tracker).

E. Circle Hough Transform

En el trabajo de Abulwafa et al. [25], Tiene como objetivo, encontrar patrones circulares del radio dado R dentro de una imagen. Cada punto de borde forma un círculo de radio R en un espacio acumulador de salida; se detecta el pico en el espacio del acumulador de salida, donde estos círculos contruibidos se superpondrán en el centro del círculo original. Se ha mejorado este método con la finalidad de reducir el tiempo de análisis computacional o de proceso, por lo tanto, se ha implementado una formulación de convoluciones aplicadas a una imagen de magnitud de borde (después de una detección de borde adecuada) [26].

F. SSD

Se presentó el modelo de detección de objetos SSD: Single Shot Multibox Detector por Wei Liu et al. en [27]. Esta técnica se denomina así porque las tareas de localización y clasificación de objetos se realizan en una sola pasada por la red, siempre en tiempo real. Multibox, porque es el nombre de una técnica de regresión de recuadros delimitadores que se integrado en el DSS. Esta técnica se inspira en el trabajo de Szegedy en [28] y consiste en una propuesta rápida de coordenadas de cuadro delimitador independientes de la clase. La detección de objetos en el modelo SSD consta de dos partes: extracción de mapas de características y aplicación de filtros convolucionales para la detección.

En resumen, las SSD utilizan una única red neuronal profunda. Se basa en una red convolucional feed-forward que produce una colección de cuadros delimitadores de tamaño fijo y puntúa la presencia de instancias de clase de objeto en estos intervalos,

seguido de un paso de supresión no máxima (análogo a YOLOv3) para las detecciones finales. Las primeras capas de la red tienen una arquitectura estándar utilizada para la clasificación de imágenes de alta calidad, y se denominó red de base. A continuación, se añade una estructura auxiliar a la red para producir detecciones, se permite extraer características a varias escalas y disminuir progresivamente el tamaño de la entrada para cada capa posterior.

G. RetinaNet

En el trabajo de Lin et al. [29], menciona que RetinaNet es uno de los mejores modelos de detección de objetos de una etapa que ha demostrado funcionar bien con objetos densos y de pequeña escala. Por esta razón, se ha convertido en un modelo popular de detección de objetos para ser utilizado con imágenes aéreas y satelitales [30].

H. VSM

En el trabajo de Macaulay y Shafiee [6], VSM es un algoritmo de aprendizaje automático supervisado que se puede aplicar tanto a enfoques de clasificación como de regresión. El algoritmo VSM identifica un “límite de decisión” o “hiperplano” para separar un conjunto de datos en dos clasificaciones distintas. El algoritmo intenta maximizar la distancia entre los puntos de datos más cercanos de las dos clases dentro del conjunto de datos. Los vectores de soporte son los puntos de datos.

I. ANN

En el trabajo de Macaulay y Shafiee [6], las redes neuronales artificiales (ANN), que pueden resolver problemas de regresión y clasificación, están modeladas en las redes neuronales del cerebro humano. Al igual que el cerebro humano que contiene miles de millones de células neuronales que están conectadas y distribuyen señales en el cerebro humano, las RNA están formadas por neuronas artificiales, llamadas unidades, agrupadas en tres capas diferentes. La primera capa se llama 'capa de entrada' que recibe datos y luego los envía a la segunda capa llamada 'capa oculta'. La capa oculta realiza cálculos matemáticos sobre los datos recibidos de la capa de entrada. La última capa es la capa de salida, que devuelve datos como salida. Las redes neuronales profundas (a veces llamadas aprendizaje profundo en la literatura académica) se refieren a redes neuronales que contienen múltiples capas.

J. KNN

En el trabajo de Macaulay y Shafiee [6], menciona que el algoritmo KNN es un algoritmo de aprendizaje automático supervisado, que se adapta mejor a los modelos de clasificación. El algoritmo realiza la estimación de la probabilidad de que un nuevo punto de datos pertenezca a un grupo en particular. Este proceso implica mirar los puntos de datos en proximidad y luego identificar qué puntos de datos tienen características similares al nuevo punto de datos. Luego, el nuevo punto de datos se asigna al grupo que tiene la mayoría de los puntos de datos con características similares cerca del nuevo

punto de datos. El algoritmo KNN es muy fácil de implementar y rápido de ejecutar. Sin embargo, KNN no clasifica muy bien los puntos de datos y la precisión del algoritmo depende de la calidad del conjunto de datos.

K. DNN

En el trabajo de Macaulay y Shafiee [6], menciona que una red neuronal profunda (DNN) es una RNA que tiene más de una capa oculta (de ahí el término "profunda") y que se entrena con grandes cantidades de datos. Cada capa oculta se compone de neuronas que asignan una función a la entrada para proporcionar una salida. Las DNN se entrenan mediante el ajuste de sus neuronas, sesgos y características de peso. Estos tipos de redes neuronales también se apoyan en diversas técnicas como el algoritmo de retro propagación y métodos de optimización como el descenso de gradiente estocástico.

Tres tipos de redes neuronales profundas son los perceptrones multicapa (MLP), las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN). Las DNN admiten sistemas de reconocimiento de voz y sistemas de traducción como Google Translate.

L. CNN

En el trabajo de Macaulay y Shafiee [6], se menciona que la red neuronal convolucional (CNN) es un tipo de ANN que detecta patrones y ayuda con el procesamiento de tareas basadas en la visión. CNN se compone de un algoritmo de unidad ML, llamado percepciones. Una CNN puede hacer predicciones analizando una imagen, verificar para identificar características y clasificar las imágenes en función de este análisis. CNN consta de múltiples capas que procesan y extraen características de los datos. Estas capas incluyen la capa convolucional, la unidad lineal rectificadora (ReLU), la capa de agrupación y la red totalmente conectada (FCN). La capa convolucional contiene filtros que realizan la operación de convolución, mientras que la capa ReLU realiza operaciones en elementos y genera un mapa de características rectificado. La capa de agrupación toma el mapa de características rectificado como entrada y realiza una operación de "reducción de muestreo" que reduce las dimensiones del mapa de características. La capa de agrupación luego convierte la salida de matriz bidimensional del mapa de características agrupadas en un vector lineal aplanándolo.

La capa FCN toma la matriz aplanada de la capa de agrupación como entrada y luego procede a clasificar e identificar las imágenes.

M. RCNN

En el trabajo de Macaulay y Shafiee [6], se menciona que los algoritmos de redes neuronales convolucionales recurrentes (RCNN) detectan y localizan objetos en una imagen. Para ello rectangulares alrededor de los objetos contenidos en una imagen, colocando una etiqueta o categorizando cada caja definida en una imagen, extrayendo características de la imagen mediante el algoritmo SVM y procesando las características mediante una CNN pre entrenada. En la imagen mediante el algoritmo SVM y, a

continuación, procesando las características con una CNN pre entrenada. La última etapa del proceso reúne las regiones separadas para obtener la imagen original con la identificación de los objetos de la imagen.

N. FASTER RCNN

Otra evolución del modelo R-CNN es el algoritmo Faster RCNN como se menciona en [31]. Faster R-CNN es un algoritmo de mejor rendimiento y más rápido que R-CNN y Fast R-CNN, porque sólo utiliza CNNs y no utiliza SVMs, y proporciona una única extracción de características de una imagen, en lugar de extracciones región por región de una imagen como R-CNN. Según la literatura, esto da como resultado redes de entrenamiento R-CNN más rápidas al menos nueve veces más rápido, con más precisión que R-CNN [32]. Sin embargo, lo que diferencia a Faster R-CNN a su predecesora Fast R-CNN es el uso de la técnica Region Proposal Network (RPN) [32].

O. Boosting Tracker

En el trabajo de Yan et al. [33], indica que Boosting tracker es uno de los primeros algoritmos de detección de objetos en el procesamiento de imágenes mediante la visión artificial. Este tracker debe tener un entrenamiento previo para el aprendizaje con ejemplos de positivos y negativos. Aparece por primera vez el término ROI (Region Of Interest), donde los píxeles dentro de la zona de interés son denominados como positivos, mientras que los que están afuera son considerados como negativos. Este rastreador se ejecuta en cada pixel adjunto a la ubicación anterior y se registra la puntuación del clasificador. A medida que se ingresan más ventanas o cuadros donde se encuentra el objeto de interés, el clasificador se actualiza con estos datos adicionales [34].

4 Análisis de resultados

La cantidad de métodos de detección de objetos que existe en la actualidad es demasiado amplia como para abordarla en su totalidad, por tal motivo se han seleccionado métodos generales, clásicos, tradicionales y basados en aprendizaje profundo, partiendo de la necesidad específica de este trabajo que es detectar un balón, es decir todo aquel método que se ha optimizado para la detección de un objeto redondo que cambia de tamaño y profundidad normalmente de color blanco y considerando que la detección debe ser en tiempo real se ha seguido el camino que se observa en la **Figura 1**, hasta conseguir una cantidad en la que es posible realizar una comparación la misma que se muestra en la **Tabla 3**.

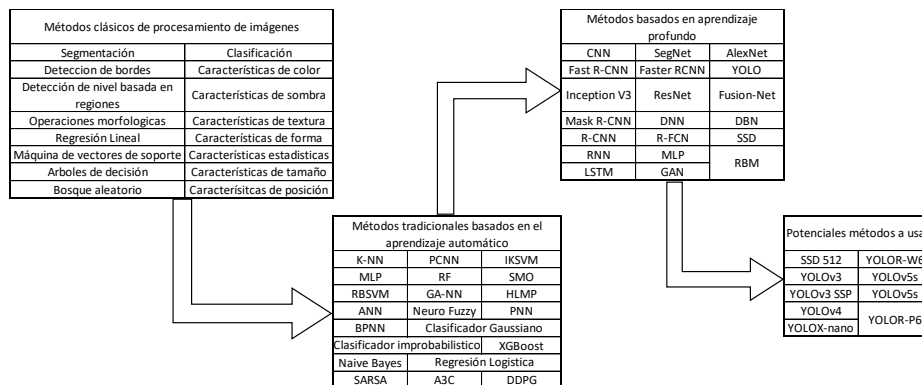


Fig. 1. Métodos de detección de objetos encontrados

Tabla 3: Métodos de detección de objetos mencionados en este documento

Técnica	Ventaja	Desventaja	Tiempo Real
YOLO	Es ideal para procesamiento de imágenes y videos en tiempo real. El sistema analiza la imagen completa. Arroja un vector con información sobre la región de imagen donde se encontró el objeto de interés, ubicación del centro de la región, distancia en ancho y alto de imagen; y clase a la que pertenece el objeto.	Pérdida de clasificación, que son las probabilidades condicionales de cada clase de objeto de interés. Pérdida de localización, mide errores entre ubicaciones y tamaños de los cuadros delimitadores.	Si
Backpropagation	Red neuronal de fácil aprendizaje, multicapa permite utilizar varios métodos de detección según necesidades de cada aplicación.	Esta red neuronal dependiendo de los datos usados para su entrenamiento es susceptible a acumular errores.	Si
KCF	Mejora la precisión y la velocidad, respecto a los anteriores. Compatible con OpenCV 3.1 y superior.	No se recupera de la oclusión total entre objetos.	Si
MIL	Tiene buenos principios para corregir errores de oclusión temporal entre objetos.	No se recupera de la oclusión total entre objetos.	Si
CHT	Ayuda a la detección de círculos en una imagen, inclusive si los círculos están ocluidos, traslapados o en ambientes ruidosos. Menor tiempo de	Pérdida de precisión en determinación de los parámetros del círculo.	No

	procesamiento y grandes cantidades de memoria.		
SSD	Con respecto a otros es más rápido y cuenta con una eficiencia superior.	Existen versiones actuales de YOLO y otros que lo superan y utilizan menor espacio de procesamiento.	SI
RetinaNet	Supera en precisión a los detectores de 2 etapas manteniendo la velocidad de los detectores de una etapa.	La velocidad a pesar de ser superior a otros detectores no consigue llegar a contemplarse como en tiempo real.	No
VSM	Funcionamiento optimizado para conjuntos de datos pequeños, el algoritmo proporciona una mayor precisión que los algoritmos contemporáneos.	Estos algoritmos no funcionan correctamente con grandes conjuntos de datos. Los tiempos de entrenamiento con grandes conjuntos de datos son altos.	No
ANN	Los algoritmos ANN se pueden aplicar tanto a problemas de clasificación como de regresión. Las redes neuronales pueden modelar datos no lineales con una gran cantidad de entradas, por ejemplo, imágenes, audio y texto.	Las ANN requieren una gran cantidad de datos para entrenar y son computacionalmente costosas. Los algoritmos de redes neuronales funcionan lentamente con las CPU tradicionales y requieren GPU.	Si
KNN	Es un algoritmo simple de entender e implementar. Es un algoritmo capaz de adaptarse a nuevos datos y no hace suposiciones sobre los datos.	Esta técnica es sensible a valores atípicos y no funciona correctamente con datos no estructurados.	Si
DNN	Es adecuada para aplicaciones que requieren una gran cantidad de datos.	Es de difícil entrenamiento, necesita una gran cantidad de datos.	Si
CNN	Los algoritmos CNN detectan características de forma independiente y automática, y son computacionalmente eficientes.	El modelo CNN no puede detectar objetos dentro de una imagen, en otras palabras, no puede decirnos dónde se encuentran los objetos. Las CNN no codifican la posición y la orientación de un objeto en su enrutamiento de datos entre capas y en los resultados o predicciones resultantes.	Si

RCNN	La técnica R-CNN es capaz de detectar objetos en imágenes, así como clasificarlas.	El proceso de varios pasos requiere mucho tiempo para entrenar la red debido a la gran cantidad de regiones (o características) que necesita clasificar para cada imagen. El modelo también requiere mucho espacio en disco y es un algoritmo computacionalmente costoso. El cuadro delimitador generado no da ninguna indicación sobre la forma del objeto.	Si
Faster RCNN	R-CNN más rápido crea una red de una sola etapa, que es más eficiente y tiene un mejor rendimiento que el proceso de múltiples etapas de R-CNN. Faster R-CNN también utiliza la técnica de agrupación de ROI para compartir datos de cómputo para el procesamiento de todos los objetos en una imagen.	El proceso de muestreo de varias regiones de una imagen da como resultado una degradación del rendimiento.	Si
Boosting Tracker	Requiere de menor tiempo de procesamiento y menor cantidad de memoria para procesamiento.	No es confiable, no existe diferencia en saber si ha fallado o no el algoritmo.	Si

En la **Tabla 3**, se pudo reducir todas las técnicas que utilizando visión artificial podrían utilizarse para detectar un balón en tiempo real utilizando un video pregrabado de un partido de fútbol, además de indicarse cuales son las ventajas que posee cada una de ellas, así como sus falencias y si cumplen con el requisito indispensable de implementarse en tiempo real, una vez presentada esta recopilación de técnicas es posible emitir un criterio de selección para la aplicación de interés.

5 Conclusiones y Trabajos Futuros

Este documento presentó a través de una extensa revisión del estado de arte, un recopilado de las técnicas y herramientas para el aprendizaje automático especializado en la visión artificial, además este estudio reveló, mediante la investigación documental, que no existe una gran variedad de documentos que reúnan las técnicas para la detección y seguimiento de objetos con sus respectivas características, mejoras, ventajas y desventajas entre cada una de ellas. También se puede concluir que los métodos de aprendizaje profundo son los más eficaces y eficientes al momento de realizar las tareas

de seguimiento y detección de objetos, debido a la gran cantidad de datos que se necesita para su entrenamiento lo que conlleva a una correcta clasificación, detección, interpretación y predicción, de los objetos de interés en las imágenes procesadas, brindando así un alto porcentaje de confiabilidad y un bajo porcentaje de errores.

Como futuro trabajo se propone utilizar una o varias de estas técnicas para la detección y seguimiento de balón mediante su acoplamiento con un sistema de posicionamiento asistido de balón para partidos de fútbol con un servidor HBBTV integrado a un guante háptico accesible para personas con discapacidad visual, permitiendo de esta manera la automatización total del mismo.

A pesar que, se tiene como prioridad la inclusión de las personas con deficiencia visual, las metodologías propuestas tienen alcance para distintas aplicaciones como, para el análisis estadístico de un partido de fútbol, lo cual sería de gran ayuda a los técnicos deportivos para mejorar las estrategias de juego, tener datos de un jugador en específico luego de un partido de fútbol y aumentar su rendimiento; otro alcance es para la detección de rostros de personas a través de cámaras de vigilancia, a esto se podría adjuntar una base de datos con los rostros de cada una de las personas de un estado o nación y sería de gran ayuda para el sistema judicial y penal de los gobiernos; o inclusive para la detección de tránsito vehicular, el cual es un problema de las grandes ciudades, y brindar posibles rutas alternativas para los conductores.

6 Referencias

- [1] L. T. Hurtado y M. A. Agudelo, «Inclusión educativa de las personas con discapacidad en Colombia,» vol. 2, n° 1, p. 11, 2014.
- [2] A. C. González, F. J. García y R. Ramírez, «Discapacidad visual,» Culcyt, vol. 2, n° 51, p. 13, 2013.
- [3] H. D. Najeeb y R. F. Ghani, «A Survey on Object Detection and Tracking in Soccer Videos,» MJPS, vol. 8, n° 1, pp. 1-13, 2021.
- [4] D. Villamarín y J. M. Menéndez, «Haptic Glove TV Device for People with Visual Impairment,» Sensors, p. 27, 2021.
- [5] J. Yuan, C. Chen, W. Yang, M. Liu, J. Xia y S. Liu, «A survey of visual analytics techniques for machine learning,» Computational Visual Media, vol. 7, n° 1, pp. 3-36, 2021.
- [6] M. Macaulay y M. Shafiee, «Machine learning techniques for robotic and autonomous inspection of mechanical systems and civil infrastructure,» Autonomous Intelligent Systems, vol. 2, n° 1, p. 25, 2022.
- [7] L. Tan, T. Huangfu, L. Wu y W. Chen, «Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification,» BMC Medical Informatics and Decision Making, p. 11, 22 November 2021.

- [8] D. J. Mendoza y J. Pérez, «Análisis de tráfico vehicular mediante visión artificial,» Repositorio Digital, Ambato, 2019.
- [9] J. Arias , M. Á. Villasís y M. G. Miranda, «El protocolo de investigación III: la población de estudio,» Revista Alergia México, vol. 63, nº 2, pp. 201-206, 2016.
- [10] M. Rashid, M. Khan, M. Alhaisoni, S.-H. Wang, S. Naqvi, A. Rehman y T. A. Saba, «Sustainable Deep Learning Framework for Object Recognition Using Multi-Layers Deep Features Fusion and Selection,» MDPI journals, 2020.
- [11] V. Londoño-Osorio, J. Marín-Pineda y E. I. Arango-Zuluaga, «Introducción a la Visión Artificial mediante Prácticas de Laboratorio Diseñadas en Matlab,» de Introducción a la Visión Artificial mediante Prácticas de Laboratorio Diseñadas en Matlab, Medellín, Colombia, 2013, pp. 591-603.
- [12] Mathworks, Inc., «Visión artificial con MATLAB,» 2022.
- [13] N. Ketkar y M. Jojo, Deep Learning with Python, Bangalore, Karnataka, India: Apress, 2021, p. 28.
- [14] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz, Y. Jia, M. Schuster, D. Mane, J. Levenberg, M. Kudlur, L. Kaiser, R. Monga, S. Moore, D. Murray, C. Olah y J. Shlens, «TensorFlow: Large-scale machine learning on heterogeneous systems,» 2015.
- [15] Digital Guide IONOS, «Keras: biblioteca de código abierto para crear redes neuronales,» Digital Guide IONOS, 08 Octubre 2020. [En línea]. Available: <https://www.ionos.es/digitalguide/online-marketing/marketing-para-motores-de-busqueda/que-es-keras/>. [Último acceso: 12 Diciembre 2022].
- [16] M. A. Fischler y R. A. Elshlager, «The Representation and Matching of Pictorial Structures,» IEEE Transactions on Computers, Vols. %1 de %2C-22, nº 1, pp. 67-92, Enero 1973.
- [17] J. Son, M. Baek, M. Cho y B. Han, «Multi-object Tracking with Quadruplet Convolutional Neural Networks,» IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [18] P. Xingang, S. Jianping, L. Ping, W. Xiaogang y T. Xiaoou, «Spatial As Deep: Spatial CNN for Traffic Scene Understanding,» Computer Vision and Pattern Recognition, 2017.
- [19] Z. Jinghua, L. Chen, K. Frank, Z. Xin, S. Changhao, L. Zihan, J. Tao, L. Hong y Q. Shouliang, «A Multiscale CNN-CRF Framework for Environmental Microorganism Image Segmentation,» National Library of Medicine, 2020.
- [20] Joseph Redmon, Santosh Divvala, Ross Girshick y Ali Farhadi, «You Only Look Once: Unified, Real-Time Object Detection,» ArXiv, 2015.
- [21] Hua Gong, Hui Li, Ke Xu y Yong Zhang, «Object Detection Based on Improved YOLOv3-tiny,» Chinese Automation Congress (CAC), 2019.

- [22] Banoth Thulasya Naik y Mohammad Farukh Hashmi, «Ball and Player Detection & Tracking in SoccerVideos Using Improved YOLOV3 Mode,» Research Square, 2021.
- [23] C. A. Rocha y J. R. Escorcía Gutierrez, «Sistema de Visión Artificial para la Detección y el Reconocimiento de Señales de Tráfico basado en Redes Neuronales,» Eighth LACCEI Latin American and Caribbean Conference for Engineering and Technology, 2010.
- [24] E. Montoya Morales, «Detección y seguimiento de personas en video,» Pontificia Universidad Javeriana, Bogotá, 2018.
- [25] Arwa E. Abulwafa, Ahmed I. Saleh, Mohamed S. Saraya y H. A. Ali, «A new ball detection strategy for enhancing the performance of ball bees based on fuzzy inference engine,» 2021.
- [26] Milad Soltany, Saeid Toosi Zadeh y H.-R. Pourreza, «Fast and Accurate Pupil Positioning Algorithm using Circular Hough,» International Conference on Computer Communication and Management, vol. 5, 2011.
- [27] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu y A. C. Berg, «SSD: Single Shot MultiBox Detector,» ArXiv, vol. 5, 2016.
- [28] E. Dumitru, S. Christian, T. Alexander y A. Dragomir, «Scalable Object Detection using Deep Neural Networks,» arXiv, 2013.
- [29] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan y Serge Belongie, «Feature Pyramid Networks for Object Detection,» ArXiv, 2016.
- [30] Desarrolladores ArcGIS, «API ArcGIS para Python, ¿Cómo funciona RetinaNet?,» Desarrolladores ArcGIS, [En línea]. Available: <https://developers.arcgis.com/python/guide/how-retinanet-works/>. [Último acceso: 12 Diciembre 2022].
- [31] Ross Girshick, Jeff Donahue, Trevor Darrell y Jitendra Malik, «Rich feature hierarchies for accurate object detection and semantic segmentation,» ArXiv, 2013.
- [32] G. Ross, «Fast R-CNN,» IEEE International Conference on Computer Vision (ICCV), 2015.
- [33] Bin Yan, Xinyu Zhang, Dong Wang, Huchuan Lu y Xiaoyun Yang, «Alpha-Refine: Boosting Tracking Performance by Precise Bounding Box Estimation,» arXiv, vol. 3, 2020.
- [34] Ricardo Mendoza, «Boosting en el modelo de aprendizaje PAC,» vol. 3, n° 3, 2013.
- [35] H. Ros Martínez, «Implementación de un sistema de visión artificial para detección de coches en tiempo real mediante Caffe2 y C++,» Universidad Politécnica de Madrid, Madrid, 2018.

TV Notifications: understanding elderly citizens' perceptions of its potential to reduce social isolation

Juliana Camargo ^[0000-0002-7537-5697], Telmo Silva ^[0000-0001-9383-7659] and Jorge Ferraz de Abreu ^[0000-0002-0492-2307]

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

Abstract. Digital exclusion is part of the list of reasons that amplify elderly social isolation and loneliness. Since they are not digital natives, like the newer generations, frequently these individuals do not have the same easiness to exchange messages or use other tools to connect them with other people. In this sense, the present work aims to identify whether the use of television to show personalized notifications, as an intermediary device, can boost interactions between the elderly and other generations since it is a familiar gadget to this audience. For this, a focus group was made with six participants, between 64 and 81 years old, who evaluated different scenarios centered on sending notifications to the television screen, suggesting actions that culminated in social interaction. Therefore, this article presents the results of this focus group. Thus, it was possible to evaluate the perception of the elderly about the potential that television must connect generations and, consequently, reduce social isolation and loneliness.

Keywords: elderly; notifications; social isolation; TV; iTV.

1 Introduction

Population ageing is one of the main current concerns of organizations such as the World Health Organization (WHO) and the United Nations (UN) [1] [2]. There are currently 1.1 billion people over 60 years of age in the world, equivalent to 13.9% of the population [1]. All countries in general have seen exponential growth in the proportion of citizens in recent years and the projection by 2030 expected are, 1 in 6 inhabitants are 60 years or older [1]. This scenario highlights the importance of investing in policies, projects and initiatives capable of inserting this public into society, ensuring autonomy and inclusion. In this sense, according to the WHO [1], technological resources, such as the internet and mobile devices, are essential for this to happen. Although they are individuals more likely to have resistance to innovation [3], democratizing and fostering access is crucial to ensure a healthier and more active life [1].

Television, in turn, plays an important role in this context because it is a device widely used by the elderly [4]. It is also the main leisure activity of this group of individuals due to accessibility and the fact that it does not involve barriers to its use [5]. The gadget is also considered a kind of company, especially for those suffering from loneliness and/or social isolation – only in Portugal, for example, 44,400 elderly people

are living in these conditions [6]. During the most critical phase of the pandemic caused by Covid-19, 16% of Portuguese individuals over 65 years old stated they spent more time in front of the TV precisely to alleviate this type of feeling [7].

This study, therefore, connects these three aspects: seniors, TV and reduction of social isolation/loneliness rates. The objective is to identify whether integrating television features that facilitate access to technological resources can make their use simpler and, consequently, reduce the rates of social isolation. For this, we sought to understand the perception of the senior individuals themselves about the theme and the potential of television to expand contact with friends and family.

To present the findings of this study, this article has been divided into four sections. The first details the methodological approach chosen to identify the opinions of the older adults. Next, the perceptions and habits of the interviewees in relation to the use of television and to the notifications are presented. The results collected from these individuals are released in the sequence – along with the analysis of the participants' discourse. Finally, the conclusions obtained and the future works carried out in the scope of this study are presented.

2 Methodological approach

Previous studies [8] [9] [10] [11] [12] showed that notifications can be mechanisms with great potential to stimulate the use of technological resources and interaction with family members, caregivers and friends. Overall, these studies have presented evidence that notifications (or reminders) have the potential to facilitate access to technology and promote interaction between older adults and individuals from other generations.

A similar conclusion was indicated in the study [19], which contains a systematic review on the subject. A thorough analysis of 29 papers involving "notifications and seniors" showed that the messages are relevant to connecting generations and promoting health-related care. The studies analyzed showed that the mechanism was important to facilitate the communication of the elderly with their family and friends. In some cases, the tool has also contributed to reducing indicators of social isolation and loneliness. These studies also supported the development of a questionnaire, which was applied to 20 Portuguese and Brazilian older adults (whose results were also presented in [19]). The conclusions obtained from the participants' responses show that, because TV is a device widely used by the older adults, using it as an intermediary of interactions (through notifications) is an effective way to promote interactions. This perception was identified mainly among the older interviewees.

Based on this previous analysis of existing studies, we sought to identify and understand the perception of senior users about the use of notifications on television. For this, a methodological approach based on the Grounded Theory was used [13].

The empirical collection was performed in two stages: i) a focus group with senior individuals for the validation of scenarios that were identified after the performance of 20 semi-structured interviews, and ii) field tests performed with a functional prototype (the results of this second stage will be presented in a future paper). For the treatment

of data, we used the qualitative approach of content analysis [14] based on the dimensions predefined in the script and on the analysis of the participants' discourse (work done from the transcription and the consented recording of the session).

The sample was chosen for convenience. The Senior University of Cacia, also in Portugal, appointed six of its students to the session.

3 Seniors' perception of the use of television notifications

The focus group about the use of television notifications was held in March 2022 at the University of Aveiro, lasting 1h36 in total, in a User Experience (uX) laboratory that simulated a real living room. Six individuals between 64 and 80 years old participated in this session, as well as two moderators and a professor from the Senior University of Cacia. The session began with the contextualization of the theme and the request for the signature of informed consent.

Then, the moderators asked each participant to make a brief presentation, saying their names, ages, television consumption habits and relationship with technology in general – we opted for an oral presentation instead of writing not to tire participants by submitting them to an online or written questionnaire. Then, three scenarios related to the use of television notifications were started, two focusing on the interaction of seniors with other people (family members, caregivers and friends) and one focused on health promotion. The representation was made from animated videos about 15 seconds long each, all simulating the use of the mechanism on television (it was taken care to develop a material with a visual language close to reality, precisely to facilitate the understanding of the public). After the display of each of them, a set of 4 to 6 open questions was made to measure the levels of acceptance about the proposals. Examples of questions: Do you find this scenario useful? Would you use this kind of functionality? Do you believe this kind of resource would help connect you with your family members? In the end, a graphical representation of all scenarios was displayed on the television screen to remind them and promote a discussion about other proposals and/or improvements in the options that were presented.

Regarding the consumption habits of television content, all participants stated that they watch television every day – only two individuals said they watch little, one of them because they have difficulty seeing. Asked about the typology of content they watch the most, they all said "tv news". Also mentioned were "documentaries" (2 times), "movies" (2 times), "entertainment programs" (1 time) and "sport" (1vez). Two participants stated that they usually do other activities while watching television, such as "playing crossword puzzles" or "making crafts" but leave the device on to have noise at home. One person also mentioned that he keeps the TV on in different rooms, as it is a "way to feel accompanied during the day".

Regarding the use of other technologies, all participants stated that they have mobile phones. Only 2 individuals have reported that they use their phones just to answer or make conventional calls. The rest of the group (4people) uses their phone daily to send messages, access social networks and make video calls. Concerning other devices, in

addition to the phone, 3 people claimed to have a computer, 1 has a tablet and the other 1 smartwatch.

4 Results and discussion

As mentioned, 3 scenarios were presented to analyze the levels of acceptance/preferences of the group of elderly with the use of notifications on television. All scenarios (presented in Fig. 1) favor social interaction between the elderly and other generations. Therefore, the objective was to analyze whether the messages displayed on television contribute, in some way, to integrate people of different ages.



Fig. 1. Scenarios presented to participants.

Scenario 1: TV call notification

Focused on the social aspect, scenario 1 presented a notification on TV with the warning that the old user was receiving a call from her granddaughter on the phone. The message suggested that she say "ok" to answer the phone call aloud. That is, the phone was configured to answer calls automatically, recognizing the activation word that was previously configured.

In consensus, the participants considered the functionality useful. Some examples of phrases spoken by the participants: "it is a way of knowing that the phone is ringing even when we are away from it, which is quite common for me" (participant 4 – P4); "it's very practical, especially for those who have a big house or do not have the phone always in hand" (P5); "it's a feature I usually use on my smartwatch and I like it a lot. It would be interesting to receive this type of notification also on television", (P4); "I often lose my son's calls because I'm doing other things and we don't talk. I think the functionality would help to avoid this problem", (P1).

In general, at first, there was difficulty in understanding whether calls would be answered through television or directly on the phone. It was necessary to intervene and explain the scenario again, emphasizing that the notification brought only a warning about the call received on the phone. However, some participants took this moment to say that they would also like to receive and make video calls through television. "So, my husband and I could talk simultaneously with our children. Without a doubt, we would talk more often a week if there was this kind of resource," commented P2. Another highlighted aspect, this time by P1, it's a possible facility in answering calls using television. "We are already used to the device, much more than with the mobile phone. I think it would be easier to use and talk more often with those who are far away," he said.

Regarding content interruptions, most of the group said it would not be intrusive or uncomfortable to receive such a notification, even if it happened during a program, football game or movie. Only one participant (P5) stressed that he would not want to answer a call if the notification was sent during a football match, especially if the communication interrupted the content.

Scenario 2: Exchanging messages with family members

In scenario 2, also focused on the social aspect, there is an example of intergenerational communication – between grandmother and granddaughter. The young woman uses her phone to send a message to the old lady saying she misses her. Content is displayed on the TV screen via a notification. By pressing the "ok" button on the remote control, the user can easily react to the message by sending a heart emoji directly to the phone of the person who initiated the contact.

In consensus, the group considered the functionality "very interesting", mainly because it can facilitate interaction with family and friends. Some examples of phrases spoken by participants: "it is the kind of thing I would use because it is very simple to interact" (P5); "not having to type to respond facilitates interaction" (P1); "it is a way of being closer to the family - and in a very simple way" (P3); "with my children living outside the country, receiving these notifications is a way of seeing that they remember me during the day. Undoubtedly, it is a way of softening the longing" (P4). The same participant made a note about receiving the notifications. According to her, it would be interesting for the messages to be preceded by audible alerts. "So, it would get our attention if we were doing other activities," he said.

In general, the group approved this type of functionality precisely because of its practicality (just press the "ok" to send a heart emoji), an indication that television can

potentiate intergenerational communication. In addition, in consensus, they said they felt "less alone" in receiving and interacting with messages simply and practically.

Scenario 3: health monitoring

In this scenario, the focus is on health monitoring, one of the most critical characteristics for individuals living alone or socially isolated. Different studies conducted in this context [15] [16] [17] [18] show that notifications on TV contribute to softening this distance between family members and caregivers, ensuring closer follow-up.

Therefore, the scenario presented consists of a system integrated into a smartwatch that detects if the heart rate is high and automatically sends a notification to the television asking if the user has already taken their regular medication. The same messages are sent to the smartphone of family members and caregivers, ensuring closer follow-up.

From the outset, P1 and P3 said that it is a "very important" feature, especially for those who have problems with forgetfulness. "Those who put the medication in boxes separated by days of the week, when receiving a warning like this, can go to the box to check if he took it correctly", (P2).

Most participants reported that they take medication daily and that a system that reminded them could help avoid forgetfulness, especially if the warnings appeared on television. They also mentioned that it is easy to forget to take the medication. P4 even uses technology to remind her. "I put an alarm clock to be remembered through my phone and smartwatch about taking an important drug," he said.

Despite recognizing the usefulness in this functionality, P5 reported that the mere fact of questioning the user regarding taking medication is something that can have adverse effects. Some examples: causing confusion and causing the elderly to ingest the drug again and not being clear enough about which medicine should be ingested at the exact time of notification. "If I've ever taken it, when I'm questioned, I can be misled. I think this feature needs some tweaking", he commented. P6 agreed that some tuning scans might be necessary but was interested and considered the functionality "very important". The group, in consensus, said that the text of the message should be as clear as possible to avoid confusion.

Finally, when asked if they would use the smartwatch to monitor their heartbeats, all participants said they "wouldn't bother to use it on a day-to-day high".

5 Conclusions and Future Work

At the end moment of the focus group, all scenarios were recalled through graphical representations so that participants could give general opinions and suggestions for new scenarios. In general, the group agreed that:

1. Scenarios 1 and 2: these are interesting possibilities to facilitate the contact of the elderly with their loved ones, thus reducing the rates of social isolation and loneliness;

2. Scenario 3: functionality is timely to promote health care and monitoring, facilitating the sharing of information with family members. However, the text used in the messages should be well-written to avoid inducing errors related to taking medications.

In a statement, the scenarios presented were well accepted by the participants, who were receptive to television notifications. In general, the group spontaneously said that the exchange of messages via television tends to be easier for them because they are accustomed to the device (besides having a larger screen, which facilitates the visualization of the contents). Consequently, because they think it is a simpler mechanism, they believe that they would maintain greater contact with their family and friends – and by different content formats, such as text messages and video calling if the TV system behaved this type of feature. This perception shows that senior citizens are open to the use of the mechanism on television and that this type of interaction has great potential to intensify intergenerational communication. This, in turn, is an effective way to mitigate and/or reduce the rates of social isolation and loneliness so recurrent among the public.

However, to test this hypothesis in a real context, a functional prototype was later developed, in partnership with an IPTV Portuguese system, which allows the sending of personalized notifications of "social" characters to the TV. Some examples of messages: i) "Your grandson makes a birthday tomorrow! Don't forget to call to congratulate me; "How about calling your son to kill the longing?" This prototype was tested for 28 days in the houses of participants who are between 65 and 81 years old and will have its results disclosed in a future paper.

References

1. Who Homepage: Ageing and Health. In: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health>, last accessed 2022/09/07.
2. UN Homepege. In: <https://www.un.org/en/global-issues/ageing>, last accessed 2022/09/07.
3. Vaportzis, E., Martin, M., Gow, A.J.: A tablet for healthy ageing: the effect of a tablet computer training intervention on cognitive abilities in older adults. *The American Journal of Geriatric Psychiatry*, vol. 25(8), pp. 841-851 (2017).
4. Rito, F.: *You & Me TV – Televisão social para idosos* (Doctoral dissertation). In.: Repository of University of Lisboa (2015).
5. Frey, B.: *Happiness and Television Viewing*. In: *Economics of Happiness*, pp. 51-54. Springer, Switzerland (2018).
6. GNR Homepage. In: <https://www.gnr.pt/comunicado.aspx?linha=4625>, last accessed 2022/09/07.
7. ICS Homepage: https://www.ics.ulisboa.pt/sites/ics.ulisboa.pt/files/2022/inquerito_praticas_culturais_2020.pdf, last accessed 2022/09/07.
8. 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.* **98**, pp. 8–50 (2017).
9. Voit, A., Weber, D., Stowell, E., Henze, N.: Caloo: an ambient pervasive smart calendar to support aging in place. In: MUM '17: Proceedings of the 16th International Conference on

- Mobile and Ubiquitous Multimedia, pp. 25–30, Association for Computing Machinery. New York, United States (2017).
10. Hong, T., Su, Y., Lee, H., Hsieh, C., Chiu, J.: VisualLink: Strengthening the connection between hearing-impaired elderly and their family. In: CHI Conference on Human Factors in Computing Systems, pp. 67–73. Association for Computing Machinery, New York, United States (2017).
 11. Valtolina, S., Hu, L.: Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. In CHIItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter (CHIItaly '21), pp. 1-5. Association for Computing Machinery, New York, United States (2021).
 12. Santana-Mancilla, C., Anido-Rifón, E.: iTVCare: a home care system for the elderly through interactive television. In: Proceedings of the Seventh Mexican Conference on Human-Computer Interaction (MexIHC'18), vol. 1, pp. 92 (2018).
 13. Glaser, B. G., Strauss, A. L.: The discovery of Grounded Theory - strategies for qualitative research. 1st edn, Routledge, London (2017).
 14. Bardin, L.: Análise de Conteúdo. Edições 70, Lisboa (2011).
 15. Hammer, S., Seiderer, A., André, E., Rist, T., Kastrinaki, S., Hondrou, C., Kollias, S.: Design of a lifestyle recommender system for the elderly. In Proceedings of the 8th ACM international conference on pervasive technologies related to assistive environments, pp. 1-8. Association for Computing Machinery, New York, United States (2015).
 16. Mainetti, L., Patrono, L., Secco, A., & Sergi, I.: An IoT-aware AAL system for elderly people. In: 2016 International Multidisciplinary Conference on Computer and Energy Science, pp. 1-6, SpliTech 2016. IEE, Split, Croatia (2016).
 17. Kotevski, A., Koceska, N., & Koceski, S: E-health monitoring system E-health monitoring system. In: Proceedings of the 6 International Conference on Applied Internet and Information Technologies, pp. 259-265. Bitola, Italy (2016).
 18. Ramljak, M.: Smart home medication reminder system. In: 2017 25th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2017, pp. 1-5. IEE, Split, Croatia (2017).
 19. Camargo, J., Silva, T. & Ferraz, J.: Connect Elderly to Other Generations Through iTV: Evaluating Notifications' Potential. In: Abásolo, M.J., Olmedo Cifuentes, G.F. (eds) Applications and Usability of Interactive TV. jAUTI 2021. Communications in Computer and Information Science, vol 1597. Springer, Cham (2022).

Tecnologías exponenciales en el futuro de la Salud y la Medicina 6P

Javier Cabo-Salvador¹, Carlos de Castro^{*2}, Verónica Cabo¹, Isabel de Castro³, José Aguilar⁵, José Miguel Ramírez², Joaquín Aguilar², Jon Arambarri², Beatriz Sainz⁴, Isabel de la Torre⁴, Johanna Caez² y Enrique García²

¹Universidad a distancia de Madrid UDIMA
jcabo@telefonica.net

² Universidad de Córdoba, CTI, Córdoba, Spain
{carlos, z82agcoj, z52camaj, jauti22, malcaloc, [egsalci-nes](mailto:egsalci-nes@uco.es)}@uco.es

³Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC)
isabel.decastro@imibic.org

⁴ETSIT. Universidad de Valladolid
{beasai, isator}@tel.uva.es

⁵ IMDEA Networks Institute
aguilarjos@gmail.com

Resumen. Las tecnologías exponenciales como la iDTV, el metaverso, la IoT, la tecnología wearable, la inteligencia artificial y la telemedicina están transformando la atención médica de manera significativa. Aunque estas tecnologías tienen un gran potencial, aún existen desafíos técnicos, éticos y legales que deben ser abordados para garantizar la calidad y la seguridad del cuidado de la salud

La convergencia de la tecnología y la salud está impulsando una transformación sin precedentes en la Medicina. La Inteligencia Computacional y la Realidad Sintética, como el Análisis de Datos y el Aprendizaje Computacional, están liderando esta revolución, junto con herramientas como la web 3.0 y los tokens no fungibles que permiten la fusión de la realidad extendida con la web mediante la tecnología blockchain. La Medicina está experimentando cambios significativos en el uso de la Realidad Sintética, como el aprendizaje médico y la formación quirúrgica con simulaciones más realistas e intuitivas. Además, la Cirugía Personalizada y de Precisión se enfoca en la exactitud de los datos, la investigación y la docencia, permitiendo el almacenamiento y procesamiento de datos y la cirugía integrada con el Big Data. El uso de robots quirúrgicos, como el Da Vinci, y la tecnología aditiva permiten planificar cirugías complejas con mayor precisión y personalización, lo que está originando un cambio de paradigma en las actuaciones terapéuticas en la nueva Cirugía Guiada.

En este artículo se realiza una revisión bibliográfica de las tecnologías exponenciales destacando la importancia de la colaboración interdisciplinaria en la implementación de estas tecnologías en el campo de la Medicina y la salud. La formación interdisciplinaria es esencial para que los profesionales de la salud puedan desarrollar habilidades en el manejo de estas tecnologías desde una perspectiva médica y tecnológica, y abordar de manera más efectiva los desafíos que plantea la Medicina del futuro, la medicina 6P (predictiva, preventiva, personalizada, de precisión, proactiva y participativa). Con este nuevo enfoque y concepto de “Medicina Personalizada de Precisión”, a diferencia de con los modelos de medicina más tradicionales, el tratamiento ya no se va a realizar de una manera igual para todos (menú cerrado), sino que se les va a poder tratar de una manera personalizada e individualizada.

Palabras clave: Telemedicina, IoT, iDTV, u-Health, Realidad Virtual, Realidad Extendida, Realidad Sintética, Metaverso, Blockchain, Data Mining, Big Data, Machine Learning, Deep Learning, Medicina 6P, mIoT, Cirugía Guiada, Biomodelos 3D, Hospital Virtual.

Summary. The convergence of technology and healthcare is driving an unprecedented transformation in Medicine. Computational Intelligence and Synthetic Reality, such as Data Analytics and Machine Learning, are leading this revolution, along with tools such as Web 3.0 and non-fungible tokens that allow the fusion of extended reality with the web through blockchain technology. Medicine is experiencing significant changes in the use of Synthetic Reality, such as medical learning and surgical training with more realistic and intuitive simulations. In addition, Personalized and Precision Surgery focuses on precision, research, and education, allowing for data storage and processing and surgery integrated with Big Data. The use of surgical robots, such as the Da Vinci, and additive technology allow for the planning of complex surgeries with greater precision and personalization, which is leading to a paradigm shift in therapeutic interventions in the new Guided Surgery. This article will highlight the importance of interdisciplinary collaboration in the implementation of these technologies in the field of Medicine and healthcare. Interdisciplinary training is essential for healthcare professionals to develop skills in handling these technologies from both a medical and technological perspective and effectively address the challenges posed by the medicine of the future, the 6P medicine (predictive, preventive, personalized, precision, proactive, and participatory). With this new approach and concept of "Personalized Precision Medicine", unlike with more traditional models of medicine, treatment will no longer be provided in the same way for everyone (closed menu), but rather individuals can be treated in a personalized and individualized way.

1 Introducción

Las tecnologías exponenciales como la iDTV (Interactive Digital Television), el metaverso, la IoT (Internet de las cosas), la tecnología wearable, la inteligencia artificial y la telemedicina están revolucionando la atención médica a nivel global. Estas tecnologías tienen un gran potencial para mejorar la eficiencia, la calidad y la accesibilidad de la atención médica, sin embargo, existen importantes desafíos técnicos, éticos y legales que deben ser abordados para garantizar la calidad y la seguridad del cuidado de la salud [36, 37].

La iDTV, una plataforma interactiva de televisión, puede ser utilizada para proporcionar información médica y de salud a los pacientes en sus hogares, así como para ofrecer programas de educación médica y de promoción de la salud. La iDTV también puede utilizarse para la monitorización remota de los pacientes y para la entrega de atención médica a distancia. Sin embargo, es necesario abordar cuestiones como la privacidad, la seguridad y la fiabilidad de los datos para garantizar la calidad y la seguridad de la atención médica proporcionada a través de esta tecnología [38, 43].

El metaverso, una tecnología emergente de realidad virtual y aumentada, tiene el potencial de transformar la atención médica, permitiendo la simulación de procedimientos médicos y la educación médica interactiva. Además, puede utilizarse para la tele-rehabilitación de pacientes y para la monitorización remota de los pacientes en tiempo real. Sin embargo, es necesario abordar los desafíos técnicos y éticos para garantizar la seguridad y la privacidad de los datos médicos del paciente [39].

La IoT, una tecnología que permite la conexión y comunicación entre dispositivos y objetos, puede ser utilizada para la monitorización y gestión remota de los pacientes, así como para la recopilación de datos de salud. Esta tecnología puede mejorar la eficiencia y la accesibilidad de la atención médica, pero es importante abordar la cuestión de la seguridad de los datos y la privacidad del paciente [40,47].

La tecnología wearable, como los relojes inteligentes y las pulseras de actividad, pueden ser utilizados para la monitorización y la gestión remota de los pacientes, así como para la prevención y el diagnóstico temprano de enfermedades. Además, la tecnología wearable puede ser utilizada en combinación con la IoT para recopilar datos de salud en tiempo real. Es importante abordar los desafíos relacionados con la precisión de los datos y la fiabilidad de los dispositivos [41,45].

La inteligencia artificial (IA) puede ser utilizada para la toma de decisiones médicas, la identificación temprana de enfermedades y la personalización de la atención médica. La IA también puede ser utilizada en combinación con otras tecnologías, como la IoT y la tecnología wearable, para recopilar y analizar datos de salud. Sin embargo, es importante abordar los desafíos éticos y legales relacionados con la privacidad de los datos y la responsabilidad en la toma de decisiones médicas [42,46].

Finalmente, la telemedicina permite la entrega de servicios de salud a distancia y ha demostrado ser efectiva en la atención primaria, la gestión de enfermedades crónicas y la atención de emergencias. Sin embargo, existen preocupaciones en torno a la privacidad y la seguridad de la información de los pacientes, la calidad de la atención y la capacitación del personal de salud. Es necesario abordar estos desafíos para maximizar los beneficios de la telemedicina en la atención médica. [48, 49,50].

U-Health y Metaverso son tecnologías emergentes que prometen revolucionar el futuro de la atención médica. U-Health hace referencia a la salud electrónica y la implementación de tecnología de la información y la comunicación en el sector de la salud, mientras que Metaverso es una extensión de la realidad virtual que permite a los usuarios interactuar con un mundo virtual en tiempo real. Esta combinación de tecnologías ofrece posibilidades infinitas, desde la educación y la formación hasta el diagnóstico y el tratamiento médico. Las innovaciones incluyen la formación quirúrgica en la realidad sintética y la atención remota en tiempo real en entornos virtuales. [1, 2, 44]

No obstante, la implementación de estas tecnologías en el campo de la salud presenta nuevos desafíos, que van desde la seguridad y la privacidad de los datos hasta la capacitación adecuada del personal de atención médica. Para asegurar la optimización de estas herramientas y cumplir con los más altos estándares de calidad y seguridad, es esencial que los profesionales de la salud trabajen en colaboración con expertos en tecnología. [3, 4]

En cuanto a la seguridad y privacidad de los datos, la U-Health y el Metaverso pueden verse afectados por ciberataques y la exposición no autorizada de datos sensibles. La protección de la privacidad y la seguridad de los pacientes debe ser una prioridad en la implementación de estas tecnologías. Además, la capacitación adecuada del personal de atención médica en el uso de estas herramientas también es fundamental para garantizar su correcta implementación. [5, 6]

Por tanto, la combinación de U-Health y Metaverso puede tener un gran impacto en la atención médica del futuro. Sin embargo, la implementación adecuada de estas tecnologías también debe abordar los desafíos relacionados con la seguridad y privacidad de los datos, así como la capacitación del personal de atención médica. La colaboración entre profesionales de la salud y expertos en tecnología será clave para asegurar el éxito de estas herramientas en la atención médica. [3, 4]

2 Los pacientes crónicos como primera causa de gasto sanitario

Durante las últimas décadas, los países industrializados han experimentado un aumento en la longevidad y la esperanza de vida. Por ejemplo, en España, la esperanza de vida ha aumentado de 69 años en 1960 a 82 años en 2020, un aumento de 13 años. Según un modelo bayesiano publicado en *The Lancet*, se espera que la esperanza de vida en los 35 países más industrializados aumente a 87 años para 2030. Sin embargo, este aumento en la longevidad no siempre va acompañado de un aumento en los años de vida ajustados por calidad (AVAC)¹.

El aumento exponencial en la longevidad ha llevado a un aumento correspondiente en los costos de atención médica, incluyendo gastos públicos, atención especializada, atención primaria, medicamentos con receta y costos de personal, así como un aumento en el PIB. [7].

La población envejecida en Europa y España ha resultado en una sociedad dominada por pacientes crónicos multi-patológicos [59].

¹ https://www.ine.es/ss/Sate-llite?L=es_ES&c=INESeccion_C&cid=1259926380048&p=1254735110672&pagename=ProductosYServicios/PYSLayout

Los procesos crónicos representan un estimado del 70-75% del gasto total en salud en países industrializados y las enfermedades crónicas son la principal causa de visitas a atención primaria y hospitalizaciones en España, con enfermedades crónicas afectando hasta el 70% de las personas mayores de 65 años. [8]

La Organización Mundial de la Salud (OMS) ha establecido como objetivo reducir la mortalidad por enfermedades crónicas en un 25% para 2025. Por lo tanto, la gestión de pacientes crónicos y la integración de servicios sociales y de salud son un gran desafío a nivel mundial. [9]

El enfoque tradicional para el manejo de pacientes agudos a través de una gestión matricial orientada a procesos ya no es válido para pacientes crónicos complejos y se necesita un cambio de paradigma. [10]

Este cambio de paradigma requiere una perspectiva inspirada en la innovación basada en la utilidad, en la que la mayoría de los datos del paciente se registran automáticamente a través de dispositivos portátiles, IoT y aplicaciones (computación ubicua), y luego se refuerzan mediante sofisticados mecanismos de inteligencia artificial analítica y nuevas ciencias como el big data y la minería de datos, aplicando la medicina del futuro a través de la U-Health y la ingeniería biomédica con IA y el "metaverso". [11]

3 Metaverso

El metaverso, como una fusión de tecnologías de realidad virtual, aumentada y sintética, se presenta como una prometedora herramienta en diversas áreas de la biomedicina, la ingeniería biomédica, la educación médica y la cirugía [12].

El potencial de estas tecnologías emergentes se extiende desde la atención médica de los pacientes, pasando por la formación en medicina, hasta la realización de intervenciones quirúrgicas. A medida que la tecnología continúa avanzando, el metaverso seguirá ofreciendo nuevas y emocionantes posibilidades en el campo de la medicina. [13,14]

El metaverso es un nuevo universo virtual que combina elementos de la realidad aumentada, la realidad extendida, la realidad virtual y la realidad física.

Esta realidad mixta o sintética permite a los usuarios interactuar entre sí de manera descentralizada y crear una experiencia inmersiva utilizando dispositivos como gafas de realidad virtual y realidad aumentada. Según un artículo publicado en la revista "Frontiers in Virtual Reality", el metaverso es una evolución de la realidad virtual y de la realidad aumentada, en la que los usuarios pueden interactuar de

manera más realista y socialmente inmersiva, lo que abre nuevas posibilidades en áreas como la educación, la salud y los negocios [14].

El metaverso, un mundo virtual al que se accede a través de la tecnología de realidad virtual, ha sido anunciado como la próxima experiencia digital clave. Está destinado a proporcionar la próxima evolución de la interacción humana después de las redes sociales y el teletrabajo.

Además, la web 3.0 y los tokens no fungibles abren nuevas posibilidades para la creación e intercambio de activos digitales. Según un artículo publicado en la revista "Blockchain in Healthcare Today", la tecnología blockchain y los tokens no fungibles pueden permitir a los pacientes controlar y compartir su información médica de manera segura y descentralizada, lo que podría mejorar la atención médica y la investigación clínica [15].

Estas tecnologías tienen un gran potencial en áreas como la biomedicina y la ingeniería biomédica, particularmente en la tele-rehabilitación y tele-educación de pacientes crónicos. Según un estudio publicado en la revista "Journal of Medical Internet Research", la tele-rehabilitación puede ser una opción viable y efectiva para mejorar la función física y la calidad de vida de los pacientes con enfermedades crónicas, como la enfermedad pulmonar obstructiva crónica (EPOC) y la enfermedad de Parkinson [16].

También pueden ser útiles en la educación médica y la cirugía, permitiendo simulaciones más realistas e intuitivas que pueden ayudar a los estudiantes a interactuar fácilmente con las estructuras anatómicas y analizar datos disponibles en tiempo real. Según un artículo publicado en la revista "Journal of Surgical Education", la simulación en realidad virtual puede mejorar la capacitación de los estudiantes de medicina y cirugía al proporcionar una experiencia práctica más realista y mejorando la comprensión de la anatomía [17].

En conclusión, el metaverso y las tecnologías asociadas tienen el potencial de transformar la manera en que interactuamos con el mundo digital y físico, especialmente en áreas como la educación, la salud y los negocios. A medida que estas tecnologías continúan evolucionando, es importante explorar su potencial y sus limitaciones para aprovechar al máximo sus beneficios [18].

Estas tecnologías están transformando la forma en que interactuamos con la tecnología y tienen un gran potencial en áreas como la biomedicina y la ingeniería biomédica. La revolución tecnológica en la medicina y la cirugía está cambiando la forma en que se enseña y se realiza la práctica clínica. La incorporación de estas tecnologías en la práctica clínica puede mejorar la eficiencia, la precisión y la calidad de la atención médica, mejorando la vida de los pacientes [19 20].

U-HEALTH y mIoT

La Cirugía Personalizada y de Precisión se enfoca en la exactitud, la investigación y la docencia, permitiendo el almacenamiento y procesamiento de datos y la cirugía integrada con el Big Data. Gracias a estas técnicas, se pueden crear modelos virtuales precisos y personalizados de órganos y tejidos, permitiendo a los cirujanos planificar y simular intervenciones quirúrgicas con mayor precisión. Además, estos modelos también pueden ayudar a predecir posibles complicaciones y evaluar el resultado de la intervención antes de llevarla a cabo [21].

La Realidad Sintética se utiliza cada vez más en el aprendizaje médico y la formación quirúrgica, permitiendo simulaciones más realistas e intuitivas que los modelos clásicos de experimentación animal, videos y e-learning [22].

Los profesionales de la salud pueden practicar y mejorar sus habilidades en un entorno virtual que les permite interactuar con los objetos y herramientas quirúrgicas de manera más realista. Esto ayuda a reducir los riesgos y los errores durante las intervenciones quirúrgicas reales y mejora la seguridad del paciente [23].

El procesamiento de datos y la cirugía integrada con el Big Data son herramientas clave en la cirugía moderna. El análisis de grandes cantidades de datos puede ayudar a los médicos a identificar patrones y tendencias que pueden ser útiles en la toma de decisiones. Por ejemplo, la evaluación de datos de pacientes anteriores puede ayudar a predecir el resultado de una intervención quirúrgica en un nuevo paciente y mejorar la precisión de la operación [24].

Además, la integración del Big Data con la cirugía permite la monitorización en tiempo real de los signos vitales del paciente y la retroalimentación durante la intervención quirúrgica, lo que mejora aún más la precisión y seguridad [25].

El uso de robots quirúrgicos como el Da Vinci y la tecnología aditiva permiten planificar cirugías complejas con mayor precisión y personalización. Los robots quirúrgicos pueden realizar movimientos precisos y delicados que los cirujanos humanos no pueden, lo que mejora la precisión y disminuye el riesgo de error humano. La tecnología aditiva, por su parte, permite la fabricación de modelos de órganos y tejidos a partir de datos de imágenes médicas, lo que permite a los cirujanos planificar y practicar intervenciones quirúrgicas personalizadas y precisas [26].

Este avance tecnológico está originando un cambio de paradigma en las actuaciones terapéuticas en la nueva Cirugía Guiada. La Cirugía Guiada se enfoca en la utilización de técnicas de imagen avanzadas y tecnología de navegación para guiar al cirujano durante la intervención quirúrgica en tiempo real. Esta técnica

permite al cirujano ver imágenes precisas de la anatomía del paciente en tiempo real, lo que ayuda a reducir el riesgo de dañar órganos y tejidos sanos durante la operación [26].

Además, la Cirugía Guiada también puede disminuir el tiempo de recuperación del paciente y mejorar los resultados finales de la intervención.

La computación ubicua (U-Health) y la inteligencia artificial, junto con el desarrollo de múltiples biosensores, permiten el registro de datos biológicos y moleculares como el ECG, la saturación de oxígeno, la presión arterial, la frecuencia cardíaca, la temperatura y los datos genómicos crudos, entre otros. Estos datos pueden transferirse digitalmente a un "hospital virtual" y desde allí, ser remitidos a un hospital físico desde la propia casa del paciente. Con la ayuda de la inteligencia artificial y el aprendizaje automático, es posible monitorizar de forma remota en tiempo real los parámetros fisiológicos y signos vitales que tradicionalmente solo podían medirse con sofisticados equipos en el ambiente hospitalario. Además, la inteligencia artificial y el aprendizaje automático pueden detectar qué pacientes están en riesgo de descompensación y quiénes pueden necesitar ser ingresados en un entorno tradicional hospitalario. [27,28,29,60,63]

La computación ubicua en Salud o u-Health, junto con el empleo de sensores biométricos y plataformas adecuadas, así como el desarrollo de historias clínicas electrónicas personalizadas y encriptadas, como la HiDi Sapiens y la realidad sintética, permite integrar todas estas herramientas para abordar el complejo campo de la medicina del futuro. Gracias a la información estructurada contenida en las historias clínicas de salud personalizadas y encriptadas, y a los datos genéticos y moleculares obtenidos a través de herramientas como la edición genética CRISPR-Cas9, el aprendizaje automático y la minería de datos, es posible diseñar estrategias tanto de prevención como diagnósticas y terapéuticas adaptadas a las características genómicas y bioquímicas de cada paciente. [30,31]

Entre las ventajas de la u-Health, destacan la posibilidad de manejar a los pacientes con enfermedades tanto agudas como crónicas desde un enfoque holístico, con toda la tecnología disponible alrededor del paciente para mejorar su bienestar desde su propia casa. Además, con los datos genómicos incorporados a las historias clínicas personalizadas y la ayuda de herramientas de ingeniería genética, es posible desde predecir los riesgos de enfermedad hasta modelar genéticamente el futuro biomédico de nuestra descendencia. [30,32]

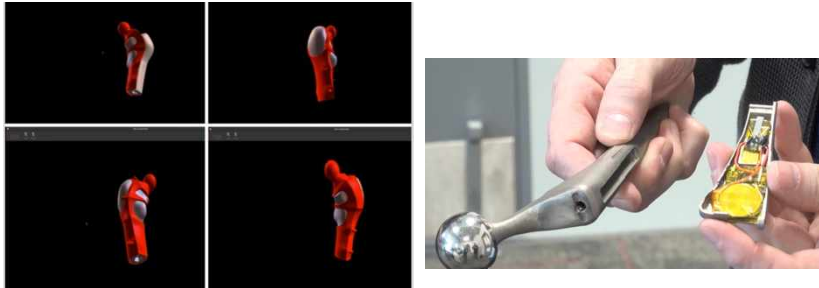


Fig.1 . Prótesis personalizada inteligente de cadera (Dr. J. Cabo)

Esta prótesis se ha obtenido mediante tecnología aditiva en 3D, incorporando chips y drogas interactivas para control del dolor. Dr. J. Cabo

La medicina personalizada de precisión, con la ayuda de la u-Health y la IA, permite predecir el riesgo de padecer enfermedades y diseñar tratamientos específicos para cada paciente. Esta "medicina a la carta" incorpora tratamientos de inmunoterapia personalizados, como los inhibidores de los puntos de control inmunitarios, y servicios avanzados como la cirugía de implantes quirúrgicos de prótesis personalizadas obtenidas mediante tecnología aditiva en 3D. Además, la realidad virtual y aumentada pueden ser utilizadas para mejorar la planificación y ejecución de los procedimientos quirúrgicos. [33,34].

La tecnología de la computación ubicua (U-Health) y la internet de las cosas médicas (mIoT) están transformando la forma en que se recopilan y se utilizan los datos de salud. Con el desarrollo de múltiples biosensores, se pueden registrar datos biológicos y moleculares en tiempo real, lo que permite la monitorización remota y la gestión de la salud del paciente. La U-Health y la mIoT también permiten la integración de datos de múltiples fuentes para crear una historia clínica personalizada y encriptada. La inteligencia artificial (IA) y el aprendizaje automático pueden analizar estos datos para detectar qué pacientes están en riesgo de descompensación y quiénes pueden necesitar ser ingresados en un entorno hospitalario tradicional. [35, 36]

La Medicina Personalizada de Precisión es un nuevo enfoque en el que los tratamientos se diseñan de manera personalizada e individualizada en base a los datos estructurados contenidos en la historia clínica de salud del paciente, su información genética y molecular. Esto permite el diseño de estrategias preventivas, diagnósticas y terapéuticas adecuadas a las características genómicas y bioquímicas de cada paciente. La U-Health y la mIoT permiten la recopilación y análisis de estos datos, lo que facilita la toma de decisiones médicas más precisas y eficaces. Además, con la ayuda de herramientas como la edición genética y el aprendizaje automático, se pueden predecir los riesgos de enfermedades y modelar genéticamente el futuro biomédico de la descendencia. [37,38]

Entre las ventajas de la U-Health y la mIoT destaca la posibilidad de manejar a los pacientes con enfermedades tanto agudas como crónicas desde un enfoque holístico, con toda la tecnología disponible alrededor del paciente para mejorar su bienestar desde su propia casa. Esto también permite una mayor participación del paciente en el cuidado de su propia salud. Además, la integración de datos genómicos en la historia clínica personalizada permite la identificación temprana de enfermedades y la prevención de enfermedades hereditarias. [35,39]

Otra área en la que la U-Health y la IA están teniendo un impacto significativo es en la cirugía de implantes quirúrgicos. Con la ayuda de la tecnología aditiva en 3D y la realidad virtual y aumentada, se pueden diseñar prótesis personalizadas que se ajusten perfectamente al paciente y mejoren su calidad de vida. Además, las prótesis pueden incorporar chips y drogas interactivas para el control del dolor. [6,7]

En resumen, la U-Health y la mIoT están transformando la forma en que se recopilan y se utilizan los datos de salud. La Medicina Personalizada de Precisión permite el diseño de tratamientos personalizados e individualizados basados en la información genética y molecular de cada paciente. La U-Health y la IA permiten la monitorización remota y la gestión de la salud del paciente, lo que facilita la toma de decisiones médicas más precisas y eficaces. Además, la integración de datos genómicos en la historia clínica personalizada permite la identificación temprana de enfermedades y la prevención de enfermedades hereditarias.

4 iDTV y gemelos digitales en u-health

La televisión interactiva (iDTV) también se presenta como un medio muy adecuado para la aplicación de la U-Health, al permitir la interacción bidireccional entre el usuario y el entorno, el acceso a contenidos multimedia, la posibilidad de personalizar la información y la adaptación a las necesidades individuales de cada usuario. Además, la iDTV también ofrece la posibilidad de implementar tecnologías de reconocimiento facial y gestual, lo que permite la creación de gemelos digitales y la personalización de los contenidos ofrecidos en base a las emociones y preferencias del usuario. Estos gemelos digitales pueden ser utilizados para crear agentes inteligentes afectivos que puedan interactuar con el usuario y proporcionar información personalizada en función de su estado emocional y físico. [51,61].

La creación de gemelos digitales también se puede llevar a cabo mediante la monitorización de las interacciones del usuario con dispositivos como los relojes inteligentes (Apple Watch, etc.) y las diademas de EEG (Muse, etc.). Estos dispositivos permiten la monitorización de la actividad cerebral, los gestos y los movimientos del usuario, lo que puede ser utilizado para crear un gemelo digital que pueda ser utilizado para la personalización de la información proporcionada

al usuario. Además, estos dispositivos también pueden ser utilizados para la monitorización de pacientes en entornos de telemedicina y la detección temprana de patologías. [52].

La interacción con agentes inteligentes también se está convirtiendo en una herramienta cada vez más utilizada en la monitorización y tratamiento de pacientes crónicos. Los agentes inteligentes pueden ser utilizados para la monitorización en tiempo real de los pacientes, la detección temprana de patologías y la intervención temprana en caso de necesidad. Además, los agentes inteligentes también pueden ser utilizados para proporcionar información personalizada y consejos a los pacientes, lo que puede mejorar el cumplimiento del tratamiento y la calidad de vida de los pacientes. La creciente popularidad de los asistentes virtuales, como GPT-3, ha demostrado el potencial de los agentes inteligentes en la monitorización y tratamiento de pacientes crónicos. [53,62].

En resumen, la U-Health se presenta como una herramienta clave en el futuro de la medicina, permitiendo la monitorización remota de pacientes, la personalización de la información proporcionada al usuario y la detección temprana de patologías. La iDTV, los dispositivos portátiles y los agentes inteligentes son herramientas clave en la implementación de la U-Health y la creación de un enfoque más personalizado y preciso de la medicina. [54].

También se está explorando la posibilidad de utilizar la inteligencia artificial para crear gemelos digitales de pacientes, que pueden ser utilizados para realizar simulaciones y predicciones de la evolución de su salud. Esto permitiría una atención más personalizada y efectiva, ya que los médicos podrían diseñar tratamientos adaptados a las necesidades específicas de cada paciente. Además, estos gemelos digitales también pueden ser utilizados para entrenar a los médicos y mejorar la precisión de los diagnósticos y tratamientos. [55].

Otra tecnología emergente en la monitorización de la salud es el uso de gemelos digitales para la creación de agentes inteligentes afectivos. Estos agentes pueden ser utilizados para interactuar con los pacientes y proporcionarles apoyo emocional y motivacional, lo que puede ayudar a mejorar su bienestar y su calidad de vida. La creación de estos agentes inteligentes se basa en la monitorización de diversas variables fisiológicas, como la frecuencia cardíaca, la presión arterial y la actividad cerebral, lo que permite obtener información sobre el estado emocional del paciente. [56].

Finalmente, cabe destacar el papel que está jugando la realidad virtual en la monitorización de la salud. La realidad virtual puede ser utilizada para la creación de entornos de entrenamiento y simulación para los profesionales de la salud, lo que permite mejorar sus habilidades y conocimientos en entornos seguros y controlados. Además, también se está utilizando la realidad virtual en la rehabilitación de pacientes, ya que permite crear entornos inmersivos y motivadores que favorecen la recuperación de lesiones. [57].

En conclusión, la combinación de diversas tecnologías emergentes como la iDTV, inteligencia artificial, los gemelos digitales, la realidad virtual, el metaverso y los agentes inteligentes está transformando la forma en que se monitoriza y gestiona la salud de los pacientes. Estas tecnologías permiten una atención más personalizada y efectiva, lo que puede mejorar el bienestar de los pacientes y la eficiencia del sistema de atención médica en su conjunto. Sin embargo, es importante tener en cuenta que estas tecnologías también presentan desafíos importantes, como la necesidad de garantizar la privacidad y seguridad de los datos de los pacientes, y la necesidad de formar a los profesionales de la salud en el uso de estas tecnologías. [58].

Conclusiones

Las tecnologías exponenciales están transformando significativamente la atención médica y están impulsando una revolución sin precedentes en la Medicina. La convergencia de la tecnología y la salud está abriendo nuevas posibilidades y oportunidades para mejorar la calidad y la seguridad del cuidado de la salud. Sin embargo, todavía existen desafíos técnicos, éticos y legales que deben ser abordados para garantizar el éxito y la efectividad de estas tecnologías en la práctica clínica.

Es esencial la colaboración interdisciplinaria y la formación interdisciplinaria para que los profesionales de la salud puedan desarrollar habilidades en el manejo de estas tecnologías y abordar de manera efectiva los desafíos que plantea la Medicina del futuro, la medicina 6P.

La Cirugía Personalizada y de Precisión es una de las áreas más prometedoras en la Medicina y la Salud, ya que permite el tratamiento personalizado e individualizado de cada paciente. La Medicina Personalizada de Precisión se enfoca en la prevención, el diagnóstico y el tratamiento personalizado de enfermedades, y utiliza tecnologías avanzadas como la Inteligencia Computacional, la Realidad Sintética y el Big Data.

En conclusión, la convergencia de tecnología y salud está impulsando una transformación sin precedentes en la medicina. Las tecnologías exponenciales como la iDTV, el metaverso, la IoT, la tecnología wearable, la inteligencia artificial y la telemedicina están transformando la atención médica de manera significativa. La Realidad Sintética y la Inteligencia Computacional están liderando esta revolución, junto con herramientas como la web 3.0 y los tokens no fungibles que permiten la fusión de la realidad extendida con la web mediante la tecnología blockchain. La Cirugía Personalizada y de Precisión se enfoca en la exactitud de los datos, la investigación y la docencia, permitiendo el almacenamiento y procesamiento de datos y la cirugía integrada con el Big Data. La medicina 6P (predictiva, preventiva, personalizada, de precisión, proactiva y participativa) representa un enfoque y concepto innovador en la atención médica, donde el tratamiento ya no

se realiza de una manera igual para todos, sino que se les puede tratar de una manera personalizada e individualizada. Aunque estas tecnologías tienen un gran potencial, todavía existen desafíos técnicos, éticos y legales que deben ser abordados para garantizar la calidad y seguridad del cuidado de la salud.

Referencias

1. Springer-Verlag: Internet of Things and Big Data for Healthcare Technologies. <https://www.springer.com/gp/book/9783030788483> (2021).
2. Xia, L., Wang, Y., Sun, X., & Yang, Y.: Virtual Reality in Healthcare: A Review. *Virtual Reality & Intelligent Hardware*, 3(4), 295-317. <https://doi.org/10.1016/j.vrih.2021.05.002>
3. Jia, J., & Liu, X.: Security and Privacy in U-Healthcare Big Data: Opportunities and Challenges. In *Big Data Analytics for Cyber-Physical Systems: Machine Learning for the Internet of Things* (pp. 265-290). Springer. https://doi.org/10.1007/978-3-030-41964-9_9 (2020).
4. Kshetri, N.: Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, 39, 80-89. <https://doi.org/10.1016/j.ijinfomgt.2017.12.005> . (2018).
5. Koufi, V., & Malamateniou, F.: The role of training and education on adopting new healthcare information systems: Lessons learned from a systematic literature review. *Journal of Biomedical Informatics*, 84, 1-12. <https://doi.org/10.1016/j.jbi.2018.07.014> (2018).
6. Imran, M., Din, I. U., Ahmad, N., Ali, G., & Nawaz, R.: Secure Authentication for U-Healthcare Using Blockchain. In *Proceedings of the 3rd International Conference on Computer Science and Education Technology (ICCSET 2021)* (pp. 424-430). Springer. https://doi.org/10.1007/978-981-33-6671-6_49 (2021). Abadía-Taira, M. B. (2019). La cronicidad, un nuevo paradigma. *Panorama social*, 29, 41-51.
7. *Health at a Glance: Europe: State of Health in the EU Cycle*. (2020). OECD. 2020
8. Muñoz-Rivas, N., Cánovas-Martínez, E., & Abadía-Taira, M. B.: Medicine of the future: wearable technology and big data. *Frontiers in Digital Health*, 2, 4. (2020).
9. United Nations, Department of Economic and Social Affairs, Population Division: *World Population Prospects 2019: Highlights (ST/ESA/SER.A/423)*. (2019).
10. Ramos, J., Soguero, C., Mora, I., Rojo, J.L., & Cabo-Salvador, J.: M-Health y su impacto en la calidad asistencial. En: J. Cabo-Salvador. (Ed.), *Gestión de la calidad en las organizaciones sanitarias*, (pp. 1009-1052). Madrid, España: Díaz de Santos (2014).
11. Musamih A.; Yaqoob I; Salah K; Jayaraman R; Al-Hammadi Y; Omar M; Ellahham S.: Metaverse in Healthcare: Applications, Challenges, and Future Directions. <https://doi.org/10.1109/MCE.2022.3223522> (2022).
12. Vargas-Cerdas, S., Rodríguez-Sánchez, J. F., & Rodríguez-Flórido, M. A.: Metaverse and Health: An Overview of the Latest Trends in Biomedicine, Biomedical Engineering, Medical Education, and Surgery. *Healthcare*, 9(2), 221. doi: 10.3390/healthcare9020221

- (2021).
13. Stojanovic, S. D., & Djordjevic-Kajan, S.: The role of the metaverse in medicine. *Vojnosanitetski pregled*, 77(1), 1-7. <https://doi.org/10.2298/vsp180915228s> (2020).
 14. *Frontiers in Virtual Reality: Metaverse for Virtual Reality Applications*. <https://www.frontiersin.org/research-topics/20201/metaverse-for-virtual-reality-applications>. (2021).
 15. *Blockchain in Healthcare Today.: The Future of Healthcare is Decentralized: How Blockchain and NFTs are Changing the Game*. [Artículo en línea]. Disponible en: <https://blockchainhealthcaredtoday.com/index.php/journal/article/view/123/319> [Último acceso: 26 de marzo de 2023].
 16. Mousavi, B., Wu, W., Schreurs, R., & Haakma, R.: Tele-rehabilitation for patients with chronic obstructive pulmonary disease: A systematic review and meta-analysis. *Journal of medical internet research*, 22(9), e23144. doi: 10.2196/23144 (2020).
 17. Chen, X., Tan, J., Xiang, L., & Li, J.: The application of virtual reality technology in medical education: a novel approach to improve surgical skills and teamwork. *Journal of Surgical Education*, 77(5), 1135-1142. doi: 10.1016/j.jsurg.2020.03.008 (2020).
 18. Liang X, Wang Q, Jiang S, Guo Y. Metaverse: A Visionary Conceptual Framework for Future Internet. *IEEE Network*.;32(2):112-117. (2018).
 19. Liao S, Duong T, Lee J, Ghanouni P.: Virtual and augmented reality in medical education and health care: A systematic review and meta-analysis. *PLoS One*.16(2):e0246409 (2021)
 20. Hu Z, Li Y, Li Q, et al. Virtual reality technology assisting surgery for upper cervical spine diseases: a systematic review from 2009 to 2019. *Eur Spine J*. 2021;30(2):479-489.
 21. Kotecha, A., & Fowler, D. L.: *Precision Surgery*. In *Medical Robotics* (pp. 225-237). Springer, Cham. [https://doi.org/10.1007/978-3-030-30712-8_17] (2020).
 22. Yang, K., & Kwak, J.: A review of virtual reality research in construction engineering and management. *Advances in Civil Engineering*, 2020. [<https://doi.org/10.1155/2020/8858769>] (2020).
 23. Cai, H., & Zheng, Y. : Medical big data analysis and processing. In *Medical Big Data and Cloud Computing* (pp. 1-17). Springer, Singapore. [https://doi.org/10.1007/978-981-15-4065-5_1] (2021).
 24. Kim, Y. T. : Robot-assisted surgery: Current status and future directions. *Journal of Medical Systems*, 43(4), 85. [<https://doi.org/10.1007/s10916-019-1215-6>] (2019).
 25. Ding, X., Chen, X., Zhao, C., & Shang, J. : Surgical navigation technology: present and future. *Computer Assisted Surgery*, 1-7. [<https://doi.org/10.1080/24699322.2021.1895572>] (2021).
 26. Chen, T. Y., Chen, Y. T., Huang, H. K., & Wu, Y. H.: Medical image processing for computer-aided diagnosis. In *Computer-Aided Diagnosis* (pp. 1-23). Springer, Singapore. [https://doi.org/10.1007/978-981-15-4891-9_1] (2020).
 27. Poon CCY, Zhang YT, Bao SD, et al. A ubiquitous healthcare solution for eldercare on demand. *Telemed J E Health*;12(5):560-565. doi:10.1089/tmj.2006.12.560 (2006).
 28. Lymberis A, Gatzoulis L. Ubiquitous health monitoring and management: opportunities and challenges. *IEEE Trans Inf Technol Biomed*;10(1):22-45. doi:10.1109/TITB.2005.856689 (2006).
 29. Wong CK, Wang Y, Yeung DY. Ubiquitous health monitoring system for elderly with diseases. *Stud Health Technol Inform*;121:220-228. PMID: 17094509. (2006).
 30. Kim, J., & Ryu, H.: The Current Status and Future Direction of Electronic Medical Records in Korea. *Healthcare Informatics Research*, 24(4), 253–261. <https://doi.org/10.4258/hir.2018.24.4.253> (2018).
 31. Eapen, Z. J., Vavalle, J. P., & Granger, C. B.: Precision medicine and personalized care for heart failure: will machine learning be the game changer?. *JACC. Heart failure*, 7(5),

- 371–380. <https://doi.org/10.1016/j.jchf.2018.12.016> (2019).
32. Ramos-Rodríguez, E.: CRISPR-Cas9 genome editing: current applications and future prospects for neurological disorders. *Frontiers in Molecular Neuroscience*, 13. <https://doi.org/10.3389/fnmol.2020.00111> (2020).
 33. Katsoulakis, E., Pappas, P., & Papalois, A.: 3D Printing and Virtual Reality: A Narrative Review of the Current State of Affairs in Surgery. *Annals of Transplantation*, 26, e931910. (2021).
 34. Shu, L., Li, X., Li, Z., Ren, X., & Wen, L. (2021). Intelligent diagnosis and precision treatment of breast cancer based on artificial intelligence and gene detection. *Future Oncology*, 17(5), 511-524.
 35. Ghosh, S., & Shin, S. Y. (2020). Ethical issues and considerations for designing and conducting wearable health research. *Journal of Biomedical Informatics*, 110, 1-10
 36. Iyawa, G. E. (2021). Applications of the internet of things (IoT) in healthcare: A review. *Heliyon*, 7(1), e05953. <https://doi.org/10.1016/j.heliyon.2020.e05953>
 37. Masood, A., Sharif, F., & Ali, S. A. (2021). IoT-based healthcare systems: A review. *Journal of Ambient Intelligence and Humanized Computing*, 12(6), 5861–5887. <https://doi.org/10.1007/s12652-021-03350-2>
 38. Kumar, A., & Singh, P. (2021). Interactive digital television (iDTV) based health services and applications: A comprehensive review. *Journal of Ambient Intelligence and Humanized Computing*, 12(8), 8637–8660. <https://doi.org/10.1007/s12652-021-03846-0>
 39. Wu, Q., Sun, M., Shen, L., & Du, X. (2021). Exploring the potential of metaverse for healthcare: A review. *Journal of Medical Systems*, 45(6), 64. <https://doi.org/10.1007/s10916-021-01781-5>
 40. Lugiai, F., Pham, Q. V., & Tantithamthavorn, C. (2021). Internet of things for healthcare: A comprehensive survey. *Journal of Ambient Intelligence and Humanized Computing*, 12(7), 7131–7166. <https://doi.org/10.1007/s12652-021-03369-6>
 41. Hernández-Orallo, J., Flórez-Revuelta, F., & Kose-Bagci, H. (2021). A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. *Sensors*, 21(1), 206. <https://doi.org/10.3390/s21010206>
 42. Hassan, M. A., Azam, N., Ahmad, A., & Amin, M. B. (2021). Artificial intelligence in healthcare: Past, present and future. *International Journal of Interactive Multimedia and Artificial Intelligence*, 7(2), 15–20. <https://doi.org/10.9781/ijimai.2021.03.002>
 43. Kumar, A., & Singh, H. (2021). Interactive digital television for healthcare: A systematic review. *Health informatics journal*, 27(2), 146-155. doi: 10.1177/1460458220961808
 44. Wu, J., Du, Y., Hu, X., & Gu, X. (2021). Application of metaverse technology in medical education and training. *Annals of translational medicine*, 9(1), 78. doi: 10.21037/atm-20-4002
 45. Lugiai, S., Riggio, R., Atzori, L., & Pilloni, V. (2021). Privacy in Internet of Things Healthcare Applications. In *EAI/Springer Innovations in Communications and Computing* (pp. 1-8). Springer. doi: 10.1007/978-3-030-78105-5_1
 46. Hernández-Orallo, E., Garcia-Zapirain, B., & García-Sáez, G. (2021). Wearable Technology for Healthcare: Monitoring and Diagnosis. *Sensors*, 21(11), 3631. doi: 10.3390/s21113631
 47. Hassan, T., Hasan, M., Abbas, H., Hussain, M., & Alghamdi, M. (2021). An IoT-Based Architecture for Personalized Healthcare Monitoring Systems. *Wireless Communications and Mobile Computing*, 2021, 1-17. doi: 10.1155/2021/6642619
 48. Chen, J., Lam, C. P., & Li, Y. (2021). Telemedicine and artificial intelligence in healthcare. In *Artificial Intelligence in Healthcare* (pp. 313-333). Springer,

- Cham.
49. Ohannessian, R., Duong, T. A., & Odone, A. (2020). Global telemedicine implementation and integration within health systems to fight the COVID-19 pandemic: a call to action. *JMIR Public Health and Surveillance*, 6(2), e18810.
 50. Sarfo, L. A., Awuah, R. B., Asiimwe, A., Ssendikadiwa, C., & Kouyate, S. (2021). Telemedicine in Sub-Saharan Africa: systematic review of applications in the delivery of health services. *African Journal of Primary Health Care & Family Medicine*, 13(1), 1-7
 51. Li, Y., Zhang, D., Jin, Y., Wang, H., & Liu, B. (2018). Affective computing based on personalized digital twins in the Internet of Things. *IEEE Transactions on Industrial Informatics*, 15(3), 1753-1762.
 52. Alharbi, A., Althobaiti, M., Aldosari, M., Alamri, A., & Alghamdi, A. (2021). Wearable devices and mobile applications in healthcare: A systematic review. *Health Informatics Journal*, 27(2), 146-156.
 53. Yang, X., Song, Y., & Wang, L. (2021). A review of intelligent agent technologies in healthcare. *International Journal of Environmental Research and Public Health*, 18(8), 3942.
 54. Covelo, C., Montero, J., & Cano, J. C. (2020). U-health: Una revisión de las tecnologías emergentes aplicadas a la salud personalizada. *RISTI-Revista Ibérica de Sistemas e Tecnologías de Informação*, (E36), 1-15.
 55. García-Saiz, D., Gómez-de la Cruz, E. J., & de Miguel-Molina, M. (2020). Digital Twins for Health: An Overview. *IEEE Access*, 8, 142594-142607.
 56. Izquierdo-Gracia, P., Torres-Sospedra, J., & Belmonte-Fernández, Ó. (2021). Affective Digital Twins: A Review of Their Applications and Challenges. *IEEE Access*, 9, 23853-23866.
 57. López-Menchero, E., Díaz-Caballero, A. J., Escolar-Reina, P., Vidaña-Márquez, E., & Hernández-Morales, M. (2020). A systematic review of the use of virtual reality in training and rehabilitation of healthcare professionals. *Journal of Clinical Medicine*, 9(10), 3255.
 58. Barnes, S.J., Brouthers, K.D., & Ghauri, P.N. (2021). The transformation of healthcare in the digital age: Challenges and opportunities. *International Journal of Information Management*, 57, 102290.
 59. Bodenheimer, T. & Berry-Millett, R. (2009). Care management of patients with complex health care needs. Princeton, NJ: Robert Wood Johnson Foundation. Busse, R., Blümel, M., Scheller-Kreinsen, D., & Zentner, A. (2009). Managing chronic disease in Europe: The initiative for sustainable healthcare financing in Europe. Retrieved from: <http://www.sustainhealthcare.org/navigation/CDM>
 60. Cabo-Salvador, J. (2017). El impacto de la u-health, inteligencia artificial robótica y nanotecnología en la medicina y el derecho. *Proceedings XXIV Congreso Nacional de Derecho Sanitario*, (pp.1-43). Madrid, España. https://www.aeds.org/img/aeds/files/documents_informacion/confmagistraljcabo-2pdf.pdf
 61. Cabo-Salvador, J., de Castro C, Cabo, V., Ramos, J., & López, M. (2017). El futuro de la IPTV integrada con la inteligencia artificial en la gestión socio sanitaria integrada. Disease Management. In: *Proceedings of the 6th Iberoamerican Conference on Disease Management: Applications and Usability for Interactive TV*. JAUTI (pp.115-129). https://www.researchgate.net/publication/362279336_EL_FUTURO_DE_LA_SALUD

[EL HOSPITAL VIRTUAL DIGITAL EN CASA Autores Javier Cabo Carlos De Castro Veronica Cabo isabel Isabel De la Torre Isabel De Castro J Miguel Ramirez](#)

62. Cabo-Salvador, J., Velarde, J.M., Cabo, V., de Castro, C., & Ramos, J. (2018). Monitoring and treatment of chronic patients through uHealth: Keys to sustainability (efficiency) and quality of care. *Sistemas & Telemática*, 16(45), 55- 83. https://www.researchgate.net/publication/327520938_Monitoring_and_treatment_of_chronic_patients_through_uHealth_Keys_to_sustainability_efficiency_and_quality_of_care
63. Cabo-Salvador, J., Ramirez-Uceda JM, Cabo, V., Ramos, J., & de Castro, C. (2018). Ubiquitous computing and its applications in the disease management in a ubiquitous city. *Journal of Computer and Communications*,6(3), 19. <https://www.scirp.org/journal/paperinformation.aspx?paperid=83132>

Digital Infrastructure |

Arquitectura extensible para el sistema de televisión interactiva basada en software libre

Ariel-Alfonso Fernández-Santana and Joaquín-Danilo
Pina-Amargós^[0000-0003-4619-849X] and Raisa
Socorro-Llanes^[0000-0002-2627-1912]

Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE)
jpina@ceis.cujae.edu.cu

Resumen La televisión digital interactiva (TVDi) permite transmitir datos que son interpretados por los receptores decodificadores y el usuario final los percibe con un determinado grado de interactividad. En la actualidad la mayoría de servicios y tecnologías utilizadas para lograr la interactividad en la televisión digital son limitados por el cobro de elevadas tarifas y su código fuente no se encuentra disponible para adecuar su funcionamiento a las condiciones locales. Dentro de las aplicaciones interactivas están los videojuegos, visitas virtuales y multimedias, las cuales contribuyen a hacer más atractivo y educativo el uso de la televisión. La presente investigación se plantea como objetivo proponer una arquitectura adaptada a las condiciones tecnológicas del país, que tribute a incorporar funcionalidades de interactividad avanzada en la Televisión Digital cubana. Para ello se realiza un análisis de las tecnologías disponibles optando por una solución de software libre que permita la creación de estas aplicaciones y su integración en un mismo sistema adecuado a la Televisión Digital cubana. Al ser la televisión un medio de difusión tan utilizado en la sociedad actual, la implementación de la propuesta fomenta el crecimiento de la soberanía tecnológica del país siendo un referente en el desarrollo de multimedias, videojuegos, visitas virtuales, u otro tipo de contenido interactivo para la Televisión Digital, ya que se crea una caracterización de los principales atributos que deben tener dichas aplicaciones.

Palabras claves: televisión digital interactiva, arquitectura de software, software libre.

1. Introducción

Se conoce como TVD (Televisión Digital) a un conjunto de tecnologías de transmisión y recepción audiovisual que emplea señales digitales en lugar de las analógicas tradicionales de la televisión [23]. Como se plantea en [2] y [21], la TVD presenta interesantes innovaciones respecto a la tradicional, tales como: varios formatos de transmisión, transmisiones simultáneas e interactividad. Basados en [4], [17] se puede afirmar que la interactividad permite a los canales de televisión ofrecer un conjunto de servicios adicionales al incorporar funciones avanzadas de comunicación, participación y servicios sociales para el desarrollo

de la informatización. Esta interactividad se clasifica en dos niveles: interactividad local e interactividad con canal de retorno [9]. La Televisión Digital interactiva (TVDi) es un escenario emergente en ascenso que brinda la posibilidad de transmitir contenidos informativos y recreativos utilizando este medio de gran penetración en la mayoría de los pueblos del mundo. Dentro de los contenidos más llamativos se encuentra la multimedia interactiva que incluye los videojuegos y las visitas virtuales. En su conjunto ofrecen una verdadera experiencia integrada combinando la televisión, el ordenador, la industria editorial y las telecomunicaciones.

Un producto multimedia interactivo permite al usuario iniciar y desarrollar un diálogo, consultar información, explorar, descubrir y adquirir nuevos conocimientos; por lo que destacan los grandes beneficios que brinda su uso en el ámbito de la educación y el entretenimiento. Se considera una tecnología que promueve la creatividad mediante los sistemas de computación. La producción y creación de sistemas virtuales reduce el derroche de recursos técnicos y económicos. Se enfoca en el uso de los recursos tecnológicos disponibles, sus avances y las herramientas multimedia para desarrollar productos interactivos, sencillos, y en los cuales utilizando diversas técnicas de diseño y creatividad se puede incluir mucho contenido informativo. Además, una interacción que exige del usuario facilita la atención, comprensión, y retención de información de una forma intuitiva y espontánea.

Dentro de los varios productos en formato multimedia que existen, en Cuba destacan tres importantes colecciones de software educativos: la colección Multisaber, la colección el Navegante y la colección Fututo [19], dirigidas a transmitir conocimiento de una forma más entretenida a los estudiantes de enseñanza primaria, secundaria y superior respectivamente. Sistemas como estos constituyen un valioso medio de enseñanza-aprendizaje, que ofrece variadas perspectivas de proyección al proceso docente-educativo. Algunos brindan un enfoque curricular y multidisciplinario por su relación con los contenidos de los programas de cada asignatura, y otros tributan a la formación de una cultura general integral. Sin embargo, no están preparados para desplegarse en el entorno masivo de la TVDi, forma en la cual se pudiese hacer llegar a los hogares independientemente de su nivel adquisitivo. En la actualidad la TVDi se utiliza principalmente para transmitir contenido multimedia que está controlado por empresas y corporaciones que responden a sus propios intereses por encima del de los usuarios. Respecto a la Televisión Cubana, existen ciertas condiciones tecnológicas y económicas que influyen notablemente cuando se refiere a implementar la interactividad. Un primer punto es que solo es posible la interactividad local debido a la ausencia de un canal de retorno, lo que hace la comunicación bidireccional inaplicable [20]. Además, las cajas decodificadoras con que se cuenta a nivel de país para el uso de la población son de bajas prestaciones.

Por otro lado, el contenido útil con respecto al tema y a las tecnologías pertinentes está disperso en Internet, lo que dificulta que los interesados en desarrollar aplicaciones en este entorno puedan asimilarlo fácilmente. Adquirir una tecnología similar de una empresa extranjera implicaría un gasto económico

importante, seguido a que no sería viable debido al bloqueo impuesto por el gobierno de los Estados Unidos hacia Cuba [11]. Al mismo tiempo, las tecnologías privativas con que están desarrolladas no permiten adaptar los contenidos existentes a los entornos de bajas prestaciones con que cuenta el país, ni integrarlos con otras soluciones.

En este artículo se describe el estado del arte referente a la interactividad en la TVD y a soluciones existentes que tributen a la coexistencia entre distintas normas de televisión. Seguidamente, para lograr la interactividad deseada se plantea utilizar un sistema adecuado al estándar HbbTV (*Hybrid Broadcast Broadband TV*) que funcione en conjunto con las soluciones de radiodifusión actuales y se presentan los resultados obtenidos en los escenarios de visitas virtuales 360 y videojuegos educativos donde fue validado. El acceso a dicho sistema será mediante un dispositivo externo como un smartphone, tablet o utilizando un televisor inteligente o caja decodificadora que tenga disponibilidad de Internet, permitiendo la comunicación cliente-servidor.

Teniendo en cuenta lo planteado anteriormente, la investigación se basa en proponer una arquitectura adaptada a las condiciones tecnológicas de Cuba, que tribute a incorporar funcionalidades de interactividad avanzada en la TVD cubana.

El desarrollo del tema que se plantea contribuiría a enaltecer el patrimonio tecnológico del país. Al ser la televisión un medio de difusión muy utilizado en la sociedad actual, su futura implementación se verá reflejado como un claro beneficio a diferentes sectores de la sociedad. Además, brindará soberanía tecnológica pudiendo convertirse en un referente en el desarrollo de multimedia, videojuegos, visitas virtuales, u otro tipo de contenido interactivo para la TVD, ya que se estará creando una caracterización de los principales atributos que deberán tener dichas aplicaciones.

2. Antecedentes

Los sistemas IBB (Integrated broadcast-broadband) han demostrado ser una solución válida para permitir a los organismos de radiodifusión llevar al usuario final nuevos tipos de contenidos emergentes (ITU-R). Dichos sistemas combinan servicios tradicionales con aplicaciones multifuncionales, similares a los servicios basados en la Web, razón por la cual disímiles radiodifusores europeos que han comenzado a aprovechar estas nuevas oportunidades que brinda. Por ejemplo, está el caso de la BBC (British Broadcasting Corporation) que puso en marcha un nuevo sistema de pruebas que ayuda a los fabricantes a verificar las funcionalidades y el rendimiento de dispositivos HbbTV para que las aplicaciones de televisión y los servicios de segunda pantalla se sincronicen con precisión con un programa o canal que se esté viendo en el televisor [13]. Por otro lado, el Grupo de Certificación HbbTV desarrolló una Aplicación de Operador HbbTV (OpApp) para permitir a los radiodifusores de televisión controlar la experiencia del usuario en dispositivos como Set Top Box (STB) y Smart TV funcionando a través de diferentes ecosistemas y dispositivos, adecuados para cualquier medio

de distribución (por ejemplo, cable, terrestre y por satélite) [15]. Es importante notar que IBB suele asociarse solamente a una norma de TVD, lo que da lugar a sistemas únicos [10]. Para superar esta limitación, varios estudios han propuesto soluciones que permiten la coexistencia entre distintas normas de televisión. En [21] se presenta un sistema de televisión híbrido que combina ISDB-T (*International System for Digital Broadcast Television*) y HbbTV, admitiendo diferentes tipos de mecanismos de sincronización entre la señal ISDB-T y las aplicaciones HbbTV asociadas. En [7] los autores presentan una red híbrida para la prestación de servicios lineales en dispositivos portables y móviles, basado en la cooperación entre el sistema de radiodifusión (broadcast) y el de banda ancha (*broadband*). En [3] se describe una plataforma en un escenario multidispositivo para la distribución de contenidos multimedia híbridos de extremo a extremo (*end-to-end*) compatible con HbbTV en el que se permite el uso de transmisión en vivo vía HTTP (*HyperText Transfer Protocol*) y protocolos de transporte en tiempo real (*Real-time Transport Protocols*) que no son soportados por el estándar HbbTV. Otro estudio [12] aprovecha la red de retorno terrestre y satelital en combinación con la conectividad a través de internet para a través de arquitecturas dinámicas de microservicios y el protocolo QUIC (*Quick UDP Internet Connections*) ampliar el alcance de las transmisiones televisivas a dispositivos como smartphones y tablets. Dichas investigaciones demuestran que es posible crear un sistema para añadir funciones de interactividad que funcione en conjunto con las soluciones de radiodifusión existentes. En [20] se presenta una nueva solución de software que demuestra las posibilidades de la TVDi en un escenario real. La solución llamada TVC+ recoge información útil disponible en Internet y la integra con los servicios de TVDi. Algunas de sus funcionalidades ya se han desplegado en varios escenarios [1], demostrando su utilidad en algunas áreas de los objetivos de Desarrollo Sostenible 2030 de la ONU (Organización de Naciones Unidas): Educación, Salud, Alimentación y Patrimonio [5]. TVC+ permite la transmisión de contenido en el estándar de TVDi que se requiera. Existen varios estándares técnicos que regulan el funcionamiento de un sistema de televisión interactiva. Basado en las investigaciones realizadas en [6] y [8] el estándar HbbTV destaca por las múltiples ventajas con respecto a las demás tecnologías, por lo que resulta ideal para el desarrollo de aplicaciones interactivas en la TVD. HbbTV se basa en un conjunto de estándares abiertos ya existentes, los cuales definen cómo se interactúa con los contenidos multimedia: OIPF-DAE (*Open Internet Protocol Television - Declarative Application Environment*), CEA (*Consumer Electronics Association*), DVB (*Digital Video Broadcasting*) y W3C (*World Wide Web Consortium*). El mismo va más allá de los estándares tradicionales, pues define sus propios para las interfaces gráficas. El estándar OIPF-DAE define APIs (*Application Programming Interface*) de JavaScript para entornos de televisión, así como establece modificaciones al lenguaje CE-HTML (*Consumer Electronics Association - HyperText Markup Language*) para la creación de interfaces gráficas. CEA define las APIs para los servicios bajo demanda, así como el acceso a redes UpnP (*Universal Plug and Play*) e Internet. El estándar DVB define la capa de transporte y señalización de los contenidos interactivos y W3C define los

estándares Web: HTML (*HyperText Markup Language*), CSS (*Cascading Style Sheets 3*), JavaScript, DOM (*Document Object Model*), entre otros para la presentación de los contenidos interactivos. Con estos estándares, HbbTV logra que contenidos de diferentes proveedores e incluso diferentes medios de transmisión sean accesibles a través de la misma interfaz y puedan ser procesados en dispositivos con bajas prestaciones. A su vez utiliza la especificación MPEG-DASH (*Dynamic Adaptive Streaming over HTTP*) con el fin de proveer soporte para el streaming adaptativo según las capacidades del cliente a través del protocolo HTTP. La transmisión de contenido con tasa de bits adaptativa es una de las características esenciales de HbbTV logrando una mejor experiencia del usuario en redes de alta saturación o de baja velocidad [16]. A varios años de publicado el estándar, numerosas organizaciones han creado soluciones para, generalmente, integrarlas con sus productos de hardware. Tal es el caso de Samsung, LG y Sony que han combinado sus sistemas de Smart TV con HbbTV. Diversos proveedores de televisión, sobre todo en Europa, han comenzado a transmitir contenidos bajo el estándar HbbTV. Actualmente suman más de 30 países donde ha sido desplegado el estándar, destacando, además del mercado europeo, Rusia, Australia y Nueva Zelanda [14]. Por otro lado, organizaciones gubernamentales, en conjunto con estos proveedores, han creado recomendaciones y especificaciones con las características que deben tener los equipos receptores y los productos de software para ser vendidos en sus países.

3. Arquitectura propuesta

La arquitectura propuesta para permitir la incorporación de interactividad en la TVD cubana se basa en el estándar HbbTV, fundamentada por todas las ventajas mencionadas anteriormente. Teniendo en cuenta que se requiere que el sistema posea un alto nivel de disponibilidad y que es substancial que permita el aumento de sus funcionalidades se decidió utilizar una arquitectura basada en microservicios. Esto contribuye a descomponer la aplicación en diferentes servicios, con el objetivo de obtener una alta disponibilidad, bajo acoplamiento, descentralización y tolerancia a fallos. Otra de las ventajas que destaca es respecto a la escalabilidad del sistema: debido a que cada microservicio funciona de manera independiente al resto, se simplifica el proceso de adición, eliminación o modificación de los módulos que componen la aplicación. Esto permite que a medida que se desarrollan los componentes que forman el sistema se vaya logrando un primer acercamiento a la implementación de esta solución, sin necesidad de ponerla en funcionamiento para todos los escenarios deseados desde el inicio. Por otro lado, permiten que la solución no esté anclada a una tecnología en particular, lo que propicia su evolución a medida que nuevas tecnologías emerjan.

Para garantizar la calidad de las aplicaciones desarrolladas implementando esta arquitectura se propone seguir los patrones de diseño API Gateway y Saga.

El patrón API Gateway crea para todos los clientes o aplicaciones externas un único punto de entrada que gestiona las solicitudes redirigiéndolas al servicio

o servicios apropiados. Esta práctica facilitar su integración con otros sistemas y permite cambiar y reorganizar los servicios sin afectar a las aplicaciones clientes ya que los aísla de conocer que la aplicación está estructurada en microservicios.

Por su parte el patrón Saga está orientado a resolver el reto de que los microservicios compartan datos y se comuniquen entre sí, debido a que cada microservicio se ejecuta de manera aislada y con una base de datos propia. Su implementación en el sistema plantea contar con un administrador de transacciones para gestionar y garantizar la integridad y consistencia de los datos entre microservicios. Cada transacción publica un evento o mensaje cuando actualiza la base de datos, acción que permite desencadenar el siguiente paso de la transacción y, si algún paso falla, poder deshacer las transacciones precedentes.

Como se mencionó HbbTV define los estándares Web para la presentación de los contenidos interactivos a través de W3C. La misión de W3C es llevar la World Wide Web a su máximo potencial mediante el desarrollo de protocolos y directrices que garanticen el crecimiento a largo plazo de la misma, elaborando especificaciones abiertas para mejorar la interoperabilidad de los productos relacionados con W3C (*World Wide Web Consortium*). Esto implica que las tecnologías utilizadas para desarrollar los módulos de interactividad del sistema deben responder a este estándar, lo cual se logra con el uso de DOM3 y lenguajes de marcado como HTML5, XML (*Extensible Markup Language*) o XHTML (*Extensible HyperText Markup Language*) asistido con CSS3 y JavaScript; mientras que para operaciones como el renderizado y cálculo gráfico en la GPU (*Graphics Processing Unit*) valida el uso de WebGPU (*Web Graphics Processing Unit*) y WebGL (*Web Graphics Library*). Igualmente se debe prestar atención a que no demanden una alta capacidad de recursos de cómputo. Con base en estos criterios se puede realizar una selección tecnológica que se adapte a la arquitectura propuesta.

Con el objetivo de obtener una retroalimentación temprana respecto a la validación de la arquitectura propuesta, este artículo se enfoca en dos posibles escenarios a los que puede ser aplicada para lograr la interactividad en la TVD cubana. En un estudio realizado en [22] durante la propagación de la COVID-19, sobre cómo las visitas virtuales 360 pueden reducir el estrés psicológico de las personas, los resultados indicaron una disminución de estrés y aumento del disfrute al participar en esta forma de realidad virtual. De forma similar en [18] se comprobó que el jugar videojuegos puede influir positivamente en reducir el estrés y la ansiedad en niños, adultos y adultos mayores, además de ser una potencial fuente de conocimientos según las características particulares del juego en cuestión. En consecuencia, se decidió que la arquitectura propuesta será validada en el escenario de las visitas virtuales 360 y los videojuegos.

Por tanto, se realizó un análisis de los marcos de trabajo fundamentales para el desarrollo de visitas virtuales 360 en un ambiente propicio para la TVD, de los que destacó Three.js para el manejo gráfico en el lado del cliente. La elección está basada en que es una biblioteca ligera de JavaScript que posee mecanismos para la optimización recursos, ventaja importante en el entorno de trabajo. Además, se mantiene actualizándose a menudo con nuevas funcionalidades, mejoras de

estabilidad y corrección de errores; lo que garantiza el soporte técnico. Para el desarrollo del resto de la aplicación del lado del cliente se utilizó React.js ya que al utilizar un DOM virtual permite un gran rendimiento y utiliza en gran medida las funcionalidades de JavaScript.

De forma similar, se realizó un análisis de las tecnologías disponibles para el desarrollo de videojuegos donde resaltó la biblioteca de JavaScript Pixi.js. La decisión se basa en que es totalmente libre, necesita pocos requerimientos para su implementación y al ser multiplataforma es compatible con los disímiles sistemas operativos de los Smart TV y STB.

En el caso de la aplicación en el lado del servidor se decidió utilizar Node.js con Nest.js fundamentado en que consume considerablemente pocos recursos del sistema comparado con otras opciones (Asp.Net o Spring), brinda una arquitectura predeterminada, utiliza JavaScript y facilita la realización de pruebas ya que permite desarrollar aplicaciones utilizando poca cantidad de código.

En las Figuras 1 y 2 se representan las propuestas para implementar el patrón API Gateway y Saga respectivamente en la aplicación de visitas virtuales 360. En el caso del patrón Saga se utiliza RabbitMQ como intermediario entre los eventos que publican los servicios al actualizar la base de datos.

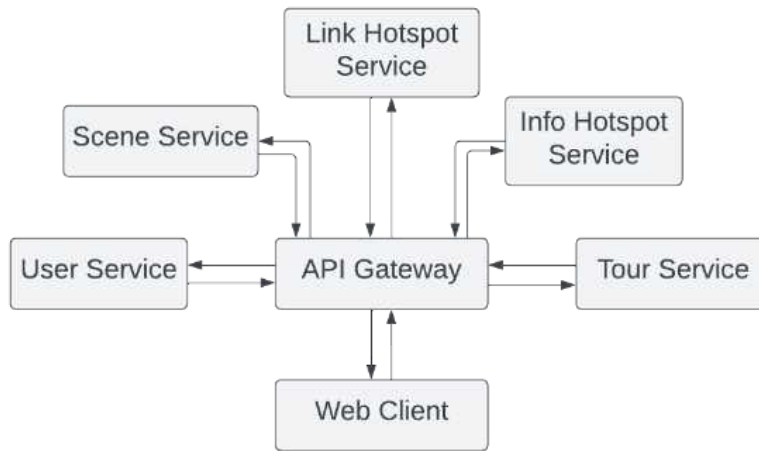


Figura 1: Patrón de diseño *API Gateway* para la aplicación de visitas virtuales 360.

En la Figura 3 se muestra el sistema representado en una arquitectura de estructuración en capas con enfoque de reutilización donde se pueden observar los diferentes paquetes que componen la solución. En la capa específica y general se muestran los paquetes particulares de la aplicación, las vistas con que el usuario interactuará, los modelos que comprenden las imágenes y recursos y la

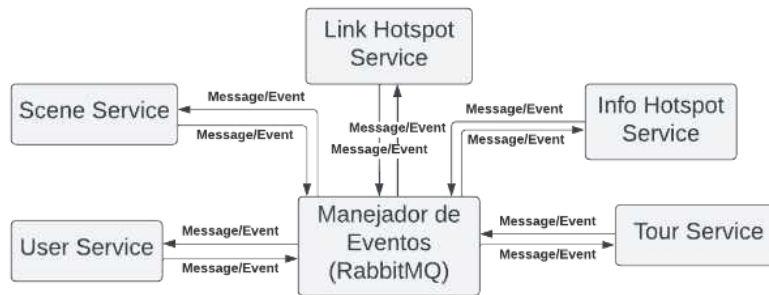


Figura 2: Patrón de diseño *Saga* para la aplicación de visitas virtuales 360.

lógica general del negocio junto a los útiles empleados; estos son los componentes menos reutilizables. En la capa intermedia se encuentran las bibliotecas y marcos de trabajo utilizados como complemento en el desarrollo del sistema, creados por terceros, y que pueden ser reutilizados para el desarrollo de aplicaciones similares. Por último, en la capa de software del sistema están contenidos los componentes propios del sistema operativo como los protocolos que gestionan el acceso a los datos y recursos.

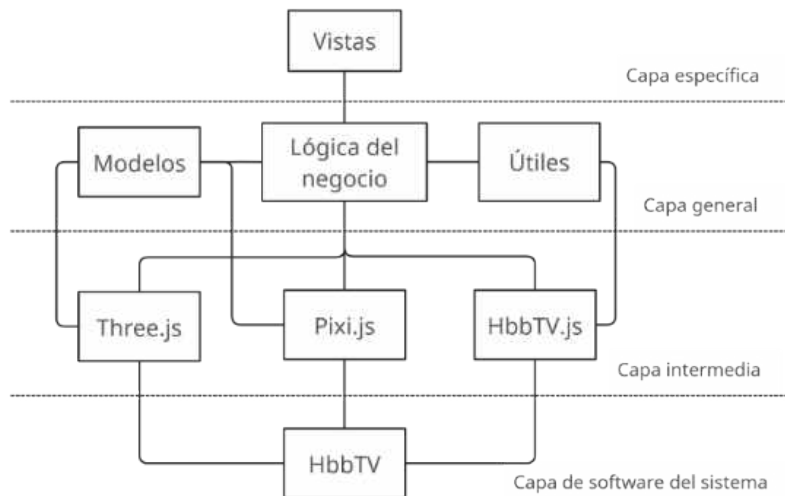


Figura 3: Arquitectura de estructuración en capas con enfoque de reutilización.

4. Análisis de resultados

Para validar la arquitectura propuesta en los escenarios señalados fueron desarrollados prototipos funcionales que permitieran comprobar su despliegue en un entorno cercano al real utilizando la tecnología seleccionada. Una vez desarrollados los prototipos se probaron en computadoras de escritorio, móviles y SBC (*Single-Board Computer*). Como navegadores Web se utilizaron Mozilla Firefox, Google Chrome y Opera.

4.1. Visitas virtuales 360

En el escenario de las visitas virtuales 360 se desarrolló un sistema que permita la conformación y visualización de estas. En la Figura 4 se muestran capturas de pantalla de su funcionamiento, tanto del editor como del visualizador. A la izquierda se observa el menú de inicio, donde se listan las visitas virtuales a las que el usuario tiene acceso y que puede visualizar, modificar o borrar. Desde esta vista también se puede crear una nueva visita virtual. A la derecha se muestra el editor y visualizador. El usuario puede conformar la visita virtual desde el editor, donde se crean las escenas y puntos de interés que la componen. Mientras que para visualizar una visita virtual se utiliza una vista que propicia la inmersión del usuario en el recorrido, a la vez que permite interactuar con los elementos del entorno.

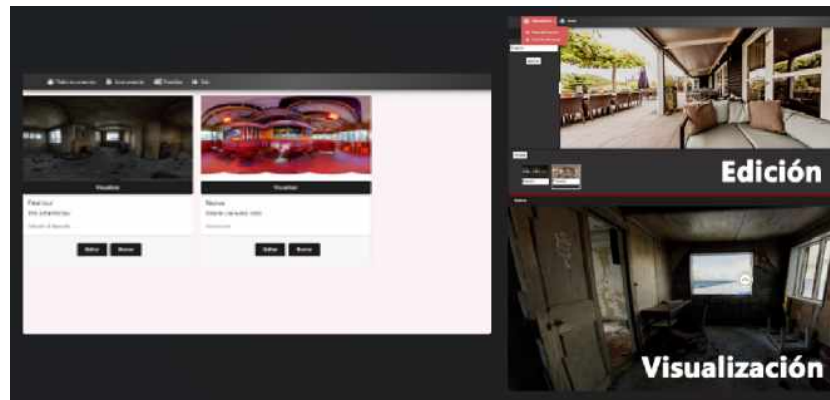


Figura 4: Sistema para la conformación y visualización de visitas virtuales 360.

4.2. Videojuego

En pos de promover el aprendizaje a través de los videojuegos se decidió que el prototipo a desarrollar fuera de lógica y conocimiento con un enfoque

de preguntas y respuestas. En función de esto, se crearon bloques de preguntas agrupadas por materia que el usuario debe responder correctamente para ganar.

En la Figura 5 se observa la vista donde se muestran las preguntas y el usuario realiza su elección. Cada partida inicia con 3 vidas representadas por corazones, cuando el jugador conteste incorrecto una pregunta perderá un corazón, que al reducirse a 0 se da por terminado el juego. Si se logran contestar todas las preguntas de la materia sin perder todos los corazones se da por vencida la materia.



Figura 5: Pantalla para la interactividad del usuario de un videojuego educativo.

Para la validación de ambos prototipos se realizaron una serie de pruebas con enfoque de caja negra, con el objetivo de verificar que cumpliera con los requerimientos planteados, lo cual arrojó resultados satisfactorios.

Con el desarrollo de estas aplicaciones siguiendo la arquitectura propuesta en esta investigación queda demostrado satisfactoriamente como puede ser implementada la interactividad avanzada en la TVD cubana en los escenarios de visitas virtuales 360 y videojuegos. Aunque se debe continuar la investigación y abordar otros escenarios como pueden ser el trabajo con segunda pantalla o videojuegos más complejos, este primer acercamiento representa un paso de avance notable en el desarrollo de este tipo de aplicaciones ya que sienta una base que los desarrolladores pueden seguir.

5. Conclusiones

El estudio del estado del arte demostró la necesidad de contar con una arquitectura adaptada a las condiciones tecnológicas del país que tribute a incorporar

funcionalidades de interactividad avanzada en la TVD cubana. Esta investigación propone dicha arquitectura y desarrolla prototipos para comprobar su correcto funcionamiento en los escenarios de visitas virtuales 360 y videojuegos.

La arquitectura propuesta está basada en el estándar HbbTV, por lo que las tecnologías a utilizar para implementar la interactividad han de adecuarse al W3C y ser de bajo impacto a nivel de recursos. Tras el análisis de las tecnologías disponibles se decidió desarrollar la aplicación cliente encargada del manejo de la interactividad con Three.js para visitas virtuales 360 y Pixi.js para videojuegos, utilizando para el lado del servidor Node.js y Nest.js.

Los escenarios validados demuestran la factibilidad de la solución propuesta y, al utilizar una arquitectura basada en microservicios, se garantiza un bajo acoplamiento entre los componentes del sistema. Esto facilita que se puedan añadir nuevas funcionalidades a la par que se reutilizan elementos ya creados, resaltando así los beneficios obtenidos. A su vez provee al sistema de la flexibilidad necesaria para sustituir los marcos de trabajo utilizados para implementar la interactividad a medida que avanza el desarrollo tecnológico y nuevas opciones emergen, sin afectar otros módulos de la aplicación; contribuyendo así a su escalabilidad y mejora continua.

Por consiguiente, el presente estudio representa un aporte para el desarrollo de la TVDi en Cuba, dado que queda sentada una base que sirve de guía y apoyo para desarrollar aplicaciones interactivas en este entorno utilizando software libre. Para dar continuidad al trabajo, se requiere validar la propuesta en escenarios reales cuando las condiciones estén creadas.

6. Agradecimientos

Esta investigación ha sido apoyada por el Fondo Fiduciario Pérez-Guerrero para la Cooperación Sur-Sur (PGTF) del Programa de las Naciones Unidas para el Desarrollo (UNDP) proyecto INT/19/K08 y el Ministerio de Ciencia, Tecnología y Medio Ambiente de Cuba (CITMA) proyecto NPN223LH006-005 de TVDi.

Referencias

1. Abásolo Guerrero, M.J., Rosado Álvarez, M.M., Silva, T., Pina, J., Socorro, R., Kulesza, R., Lemos de Souza Filho, G., De Giusti, A.E., Naiouf, M., Pesado, P.M.: 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 Transactions on Broadcasting* **64**(3), 721–746 (2017)

4. Dávila Sacoto, M.A.: Diseño de una plataforma de software para Televisión Digital interactiva de un canal de deportes utilizando GINGA-NCL LUA. B.S. thesis (2012)
5. Desa, U., et al.: Transforming our world: The 2030 agenda for sustainable development (2016)
6. Eslava Arce, I.: Estudio del estándar de televisión digital interactiva hbbtv e implementación de aplicación final (2014)
7. Fam, P.A., Paquelet, S., Crussière, M., Hélar, J.F., Brétilon, P.: Analytical derivation and optimization of a hybrid unicast-broadcast network for linear services. *IEEE Transactions on Broadcasting* **62**(4), 890–902 (2016)
8. Frómata-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)
9. García-Crespo, A., Ruíz-Mezcua, B., González-Carrasco, I., Cuadrado, J.L.L.: Servicios interactivos y accesibilidad en la televisión digital, una oportunidad para reducir la brecha digital. *Rev. Iberoam. de Tecnol. del Aprendiz.* **7**(2), 86–93 (2012)
10. Gavril, C., Popescu, V., Fadda, M., Anedda, M., Murrioni, M.: On the suitability of hbbtv for unified smart home experience. *IEEE Transactions on Broadcasting* **67**(1), 253–262 (2020)
11. Ginsburg, M.: Economic and media war against socialist societies. *International Journal of Cuban Studies* **14**(2), 272–308 (2022)
12. Hammershøj, A., Nowak, A., Hansen, J.K., Stefanović, Č.: 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)
13. Hammond, M.: Release of hbbtv/dvb companion synchronisation tools and streams. BBC R&D. Retrieved from <https://www.bbc.co.uk/rd/blog/2017-05-second-screen-streams-tools-companion> (May 2017)
14. HbbTV: Hbbtv deployment countries. Retrieved from <https://www.hbbtv.org/deployments/> (2022)
15. HbbTV-Certification-Group: The benefit of hbbtvopapp for operators and vertical models - the open standard for a unified tv experience across different platforms. Retrieved from <https://www.hbbtv.org/wp-content/uploads/2018/09/HbbTV-MG-00632-003-WhitepaperHbbTVOpApp-v1.1.pdf> (September 2018)
16. Jakšić, B., Milošević, I., Petrović, M., Ilić, S., Bojanić, S., Vasić, S.: Characteristics of hybrid broadcast broadband television (hbbtv). *Bulletin of Natural Sciences Research* **7**(1) (2017)
17. 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)
18. Pallavicini, F., Pepe, A., Mantovani, F.: Commercial off-the-shelf video games for reducing stress and anxiety: systematic review. *JMIR mental health* **8**(8), e28150 (2021)
19. 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)
20. 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 Informáticas* **12**, 158–172 (2018)

21. Sotelo, R., Joskowicz, J., Rondán, N.: An integrated broadcast-broadband system that merges isdb-t with hbbtv 2.0. *IEEE Transactions on Broadcasting* **64**(3), 709–720 (2018)
22. 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. *Technology in Society* **64**, 101514 (2021)
23. Zárate-Torres, D.J.: Diseño e implementación de una aplicación interactiva para educación a distancia: T-learning a través de un canal tdt universitario y un canal iptv en la ciudad de lima (2012)

Analysis of the value and technological acceptance of OTT platforms in Quito, Ecuador

Carina Haro Granizo¹, and Gonzalo Olmedo¹

¹ Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador
WiCOM-Energy Research Group
[caharo4, gfolmedo]@espe.edu.ec

Abstract. The appearance of over-the-top (OTT) platforms in the early 21st century replaced the physical distribution of multimedia content through videos. Once popularized and in a disruptive stage, the technological capacity to distribute more content increases demand for this service worldwide.

In this article, an investigation is carried out through a demonstration of convenience in Quito, Ecuador, to analyze the value and technological acceptance of these digital content services, determining the expected factors and those that contribute to the continuity of its use. A statistical data analysis methodology is used, where the influence of the variables determining how much the content is enjoyed, ease of use, technical risk, cost risk, and security risk that emerges on the perceived value is evaluated. A structural equations model is proposed for the analysis of hypotheses and the similarity of variables. In addition, the OTT platforms that are currently used and that have greater acceptance concerning the segment of the population are identified.

Keywords: OTT, SEM, SPSS, AMOS, satisfaction, acceptance, risk.

1 Introduction

Digital media platforms, called over-the-top (OTT), have experienced high growth globally due to the variety of content and quality of experience that can be provided based on the new Internet infrastructures [1][2]. It is customary to hear about OTT applications that offer audio and video content, such as Netflix, Disney Plus, Amazon Prime, and Apple TV +, among others, through subscriptions, which are part of the global market. The change from television by cable or satellite to OTT applications leads to a competitive process of acquiring customers and retaining them [3]. The same goes for OTT platforms due to increasing competition.

In [4], it is shown that OTT platforms will register an annual growth of 29% between the years 2021 and 2028 worldwide. In Latin America, in 2021, the percentage of households with OTT was 26%, projected to reach 35% in 2026 [5].

Some studies analyze customer satisfaction levels on OTT platforms and the tendency to stay with the service. These investigations are available in the literature for countries such as India [6] and Korea [7]. In the present investigation, it is desired to replicate the Model of evaluation of the perceived value and the intention to continue

2

using the service within Latin America, taking the city of Quito in Ecuador as a scenario. Additionally, the methodology will be based on a survey designed from the constructs and questions validated in the literature, and Structural Equation Models will be used to evaluate the results.

2 Methodology

The hypotheses related to the perceived value and the intention of use will be evaluated, applying a quantitative scientific methodology through an online survey supported by a complete structural model.

2.1 Online survey design

The number of households with OTT services for the year 2020 was 21% in Ecuador, and the trend of the mean value indicates that by the year 2022, the mean number of households with OTT services will be 28% in Latin America [5]. Based on the last census of Ecuador, from 2010, in the Quito canton, the focus of our studies was a total of 2,239,191 inhabitants [8]. With an average of 4 members per family, the total number of possible households is 559,798. It is obtained that a minimum number of surveys $n=255$ are required based on Cochran's formula for a finite population, presented below [9]:

$$n = \frac{p(1-p)}{\frac{e^2}{z^2} + \frac{p(1-p)}{N}}, \tag{1}$$

where:

n = sample size

N = population size (559,798 households).

e = acceptable sampling error (5%).

p = the population proportions (21% of households had the OTT service by 2020).

z = z value at reliability level 95% or significance level 0.05, $z = 1.96$.

A survey was created based on the constructs and questions proposed in the literature [7], keeping the questions most significant after a confirmatory factor analysis, as shown in Table 1. For the analysis, 448 surveys were carried out in Quito, with people in their 98,7% between 20 and 55 years old, according to what is presented in Figure 1.

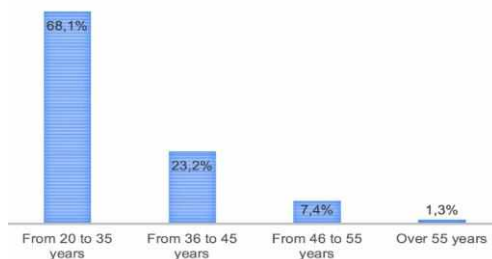


Fig. 1. Age of the population that was interviewed - 448 samples in Quito.

Table 1. Constructs, codes and items for the questionnaire.

Construct		Code	Item
C1	Enjoyment	EN1	Content from platforms helps me relax and have fun in my free time.
		EN2	The use of platforms contributes to my quality of life.
C2	Ease of Use	EU1	The management and selection of content on platforms are intuitive.
		EU2	I know how to view or change my personal information on platforms.
		EU3	I know how many devices I can use simultaneously on platforms.
C3	Usefulness	US1	I can see all the content I want.
		US2	I want to have a partial courtesy view before buying a platform.
		US3	I find the content in the language of my preference.
		US4	It is helpful to see the score in each of the contents.
C4	Technicality	TE1	Multiple devices.
		TE2	Compatibility between devices and use of platforms.
		TE3	Internet connection quality.
C5	Perceive Fee	PF1	Subscription fees.
		PF2	Satisfaction is based on price.
		PF3	Acquisition of additional equipment to enjoy the content.
C6	Personal Security Risk	SE1	Possibility of my personal information being hacked.
		SE2	My personal information will be affected by sharing accounts.
		SE3	Prevent someone from using my account on the platforms.
		SE4	Security of my credit card on the platform.
C7	Perceived Value	PEV1	I share content information with my family and friends.
		PEV2	Content type
C8	Intension to Use	IU1	The platform adapts to my preferences.
		IU2	Overall, I am satisfied with the platform I am using.

2.2 Research Model

The research model proposed in [7] was used to analyze OTT user satisfaction and the intention to continue using the service in the population of Quito, which is based on a Value-Based Adoption Model (VAM) that classifies the benefits (utility and enjoyment) and sacrifice (technological features and perceived cost) as main factors of perceived value and analyzed intention to use [10].

The proposed model is presented in Figure 2, where six exogenous variables can be observed that represent the antecedent variables of the positive factors (enjoyment, convenience, and usefulness) and negative factors (technical characteristics, cost risk, and security risk) that affect the endogenous variables of perceived value in relation to the satisfaction of use and the intention of continuous use. The exogenous variables are, at the same time, latent structures that are supported by observable variables through the results of the surveys. In this investigation for the analysis, the Structural Equation Model (SEM) was created through the SPSS and AMOS tools in version 24 [6].

4

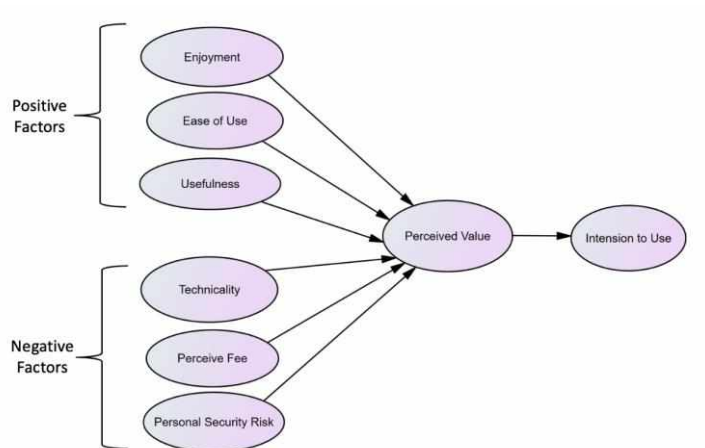


Fig. 2. OTT platform Research Model.

3 Results

The structural equation model that studies the causal relationships between endogenous and exogenous variables was analyzed using the SPSS-AMOS tool through the maximum likelihood technique. The covariance between the exogenous variables was also obtained, and the error correction variables required to compensate for the equations were included. The results were initially validated using the confirmatory analysis technique to validate the most representative questions.

The adjustment of the model was validated based on the results obtained in the indices presented in Table 2, where all are accepted since they are within the limits defined in the literature.

The estimates of maximum likelihood and the covariances of the model variables are presented in Tables 3 and 4 with an estimate of the standard error of the covariance, the critical ratio, and an approximate two-tailed p value for testing the null hypothesis respectively, the analysis of which will be presented in the discussion section. The graphic result of all the estimations is shown in Figure 3.

Table 2. Model Fit Summary.

Fit parameters		Obtained value	Condition to accept	Decision	Reference
Chi-square value	CMIN/DF	2.057	≤ 3	Acceptable	[11]
Root Mean Square Error	RMSEA	0.049	≤ 0.05	Excellent	[12]
Baseline Comparisons	CFI	0.799	Closed to 1 show a very good	Acceptable	[13]

Table 3. Maximum Likelihood Estimates.

Maximum Likelihood Estimates			Estimate	S.E.	C.R.	P
Perceived Value	<---	Enjoyment	0,867	0,769	1,128	0,259
Perceived Value	<---	Ease of Use	0,239	0,321	0,746	0,456
Perceived Value	<---	Perceive Fee	-1,001	0,448	-2,235	0,025
Perceived Value	<---	Personal Security Risk	-0,146	0,257	-0,57	0,569
Perceived Value	<---	Usefulness	-1,363	1,175	-1,161	0,246
Perceived Value	<---	Technicality	3,481	5,146	0,676	0,499
Intension to Use	<---	Perceived Value	0,629	0,124	5,084	***
EN2	<---	Enjoyment	1,261	0,205	6,144	***
EU2	<---	Ease of Use	2,092	0,435	4,811	***
EU3	<---	Ease of Use	1,895	0,393	4,825	***
US2	<---	Usefulness	0,697	0,102	6,859	***
US1	<---	Usefulness	0,978	0,214	4,575	***
TE2	<---	Technicality	3,641	2,05	1,776	0,076
TE1	<---	Technicality	3,072	1,709	1,798	0,072
US4	<---	Usefulness	1,125	0,186	6,046	***
PF2	<---	Perceive Fee	-4,443	1,098	-4,045	***
PF3	<---	Perceive Fee	-0,61	0,24	-2,54	0,011
SE3	<---	Personal Security Risk	0,579	0,202	2,866	0,004
SE2	<---	Personal Security Risk	-1,371	0,316	-4,336	***
SE1	<---	Personal Security Risk	-1,456	0,342	-4,262	***
IU2	<---	Intension to Use	2,242	0,403	5,558	***
PV2	<---	Perceived Value	-0,7	0,327	-2,141	0,032

*** less than 0,001

Table 4. Covariances.

Covariances			Estimate	S.E.	C.R.	P
Enjoyment	<-->	Ease of Use	0,007	0,002	2,847	0,004
Enjoyment	<-->	Perceive Fee	-0,01	0,003	-3,419	***
Enjoyment	<-->	Usefulness	0,016	0,003	5,852	***
Enjoyment	<-->	Technicality	0,005	0,003	1,775	0,076
Personal Security Risk	<-->	Enjoyment	0,004	0,003	1,421	0,155
Ease of Use	<-->	Usefulness	0,012	0,003	3,719	***
Ease of Use	<-->	Perceive Fee	-0,01	0,003	-2,923	0,003
Ease of Use	<-->	Technicality	0,003	0,002	1,647	0,1
Personal Security Risk	<-->	Ease of Use	0,007	0,003	2,279	0,023
Usefulness	<-->	Technicality	0,006	0,003	1,766	0,077
Usefulness	<-->	Perceive Fee	-0,012	0,003	-3,427	***
Personal Security Risk	<-->	Usefulness	-0,001	0,002	-0,285	0,775
Technicality	<-->	Perceive Fee	-0,003	0,002	-1,598	0,11
Personal Security Risk	<-->	Technicality	-0,001	0,001	-1,031	0,302
Personal Security Risk	<-->	Perceive Fee	-0,005	0,003	-2,019	0,044

*** less than 0,001

6

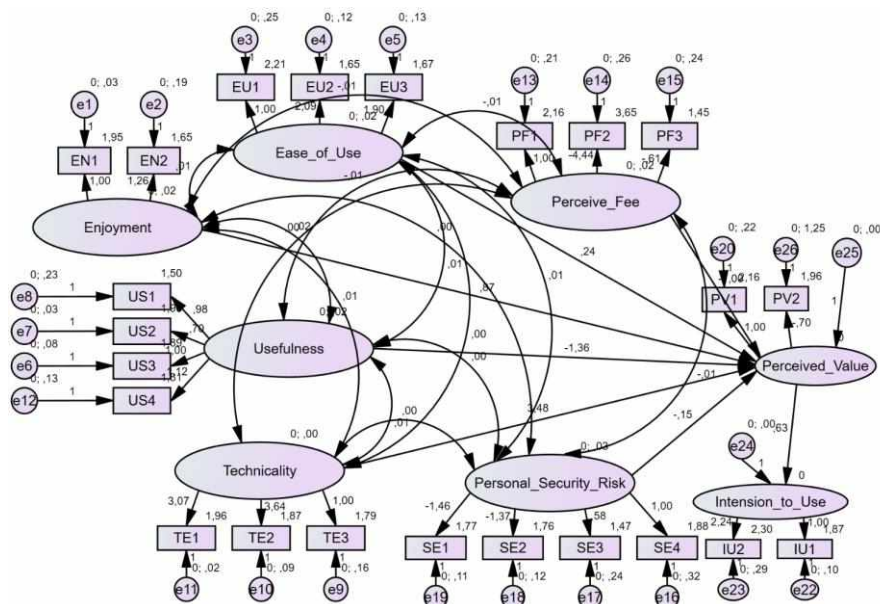


Fig. 3. Structural Equation Model.

Additionally, the most used OTT applications in Quito were registered, where Netflix is positioned in first place with 85.9%, according to the results shown in Figure 3.

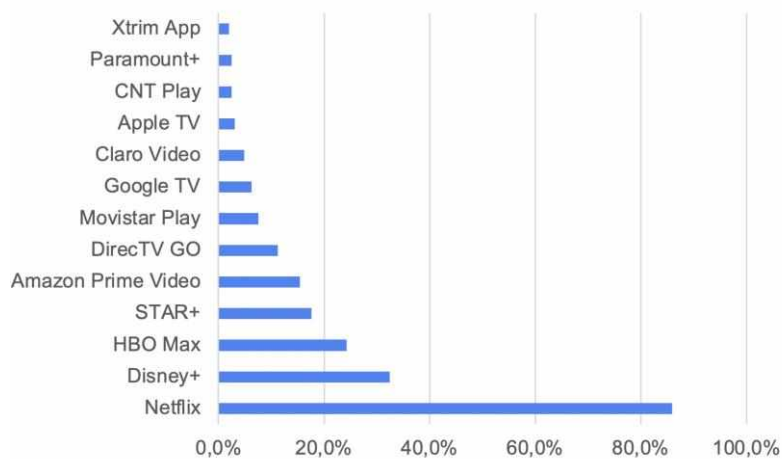


Fig. 4. Percentage of users for OTT platforms in Quito.

The survey also showed that TV is the most used device to view OTT content, with 52% based on the results presented in Figure 5a, and that people prefer to watch high-resolution content, 4K, or HD, as shown by the results in Figure 5b.

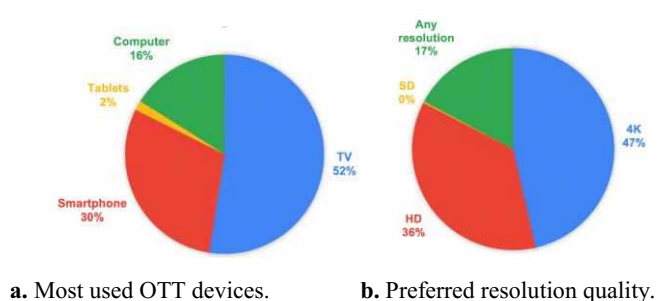


Fig. 5. Devices and preferred quality to watch OTT content.

Regarding content's OTT preference, the most recommended among users were series, followed by movies, as shown in Figure 6.

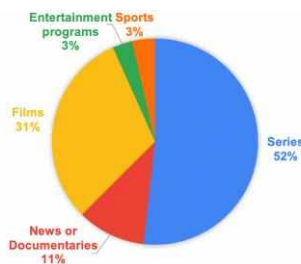


Fig. 6. Most recommended content among OTT users.

4 Discussion

From the results obtained in the maximum likelihood estimation presented in Table 3, the significant results in the first place prove the hypothesis that the perceived value significantly depends on the intention to continue using the OTT service. Additionally, validate the endogenous variables according to their probability of significance.

Regarding the perceived value, the three most significant latent structures presented by the results of the surveys in the city of Quito in order of priority according to the results are: first, with a high degree of significance, is the service fee that depends more significantly on satisfaction based on price (PF2). Secondly, the OTT platforms are usable, the fact of being able to see the OTT content when the user wants (US1), being able to test the service before contracting it (US2) and knowing the ratings of the content (US4). In third place, but very close to the second, is, with a high significance, entertainment support to improve their quality of life (EN2). The other three latent structures are not significant. However, the questions associated with each of them are significant and represent a valuable tool for hypothesis analysis.

Based on Table 4, ease of use, perceive fee, usefulness, and enjoyment are significantly related. Technology has a more significant relationship with usability, and the perceive fee is additionally significantly related to security risk.

5 Conclusions

In this article, an analysis of the value perceived by the user of OTT platforms in the city of Quito and the intention to continue with the service was carried out. The study began with building a survey based on latent structures related to constructs with questions already validated in the literature in a structured model for research. Thus, six exogenous latent variables structured by the Value-Based Adoption (VAM) model were analyzed, which classifies positive factors (enjoyment, ease of use and usefulness) and negative factors (technicality, perceive fee and personal security risk) as the main factors of perceived value and analyzes intention of use. A structural equations model is proposed that is processed in SPSS-AMOS using a maximum likelihood and covariance analysis for the analysis. This model was successfully validated in terms of fit parameters.

The results showed that the value perceived significantly affects the intention to continue using the platforms. The exogenous structure with the most likely to be significant is related to cost, usability, and enjoyment.

Additionally, results were shown where the most used OTT applications in the city of Quito are presented, as well as the preference of the technological equipment and resolution that users have the highest preference for, as well as the contents that have the highest recommendation among users.

References

1. Chang, P.-C., Chang, H.-Y., 2020. Exploring the factors influencing continuance usage of over-the-top services: the interactivity, consumption value, and satisfaction perspectives. *Int. J. Technol. Hum. Interact.* 16 (4), 118–138. <https://doi.org/10.4018/IJTHI.2020100108>.
2. Sudhir, A., Rao, P., 2021. How technology is shaping the future of streaming services in India. *Financ. Express*. January 11. <https://www.financialexpress.com/brandwagon/how-technology-is-shaping-the-future-of-streaming-services-in-india/2162729>.
3. Fudurić, M., Malthouse, E.C., Lee, M.H., 2020. Understanding the drivers of cable TV cord shaving with big data. *J. Media Bus. Stud.* 17 (2), 172–189. <https://doi.org/10.1080/16522354.2019.1701363>.
4. PWC, 2021. Global entertainment & media outlook 2021–2025. [Industry Growth Projections]. PricewaterhouseCoopers LLP 12. Outlook. <https://www.pwc.com/gx/en/entertainment-media/outlook-2021/perspectives-2021-2025.pdf>.
5. Jung, J., Melguizo A.: Is your netflix a substitute for your telefunken? Evidence on the dynamics of traditional pay TV and OTT in Latin America, *Telecommunications Policy*, ISSN 0308-5961 (2022).

6. Debarun Chakraborty, Mujahid Siddiqui, Aaliyah Siddiqui, Justin Paul, Ganesh Dash, Francesca Dal Mas, Watching is valuable: Consumer views – Content consumption on OTT platforms, *Journal of Retailing and Consumer Services*, Volume 70, 2023, ISSN 0969-6989.
7. S. -Y. Yoon and J. -B. Kim, "A Study on User Satisfaction and Intention to Continue Use of OTT Platform Digital Content Provision Service," 2022 IEEE/ACIS 7th International Conference on Big Data, Cloud Computing, and Data Science (BCD), 2022, pp. 290-296, doi: 10.1109/BCD54882.2022.9900797.
8. INEC, Census of population and housing, <https://www.ecuadorencifras.gob.ec/censo-de-poblacion-y-vivienda/>, last accessed 2022/11/04.
9. Uakarn, C., Chaokromthong, K., Sintao, N.: Sample Size Estimation using Yamane and Cochran and Krejcie and Morgan and Green Formulas and Cohen Statistical Power Analysis by G*Power and Comparisons, *Apehit International Journal*, 76-88 (2021).
10. Ji-eun Yoon, "A Study on User Perception and Use Intention of Digital Content Characteristics," Kyonggi University Graduate School, Ph.D. thesis. 2006.
11. Kline, R. B, Principles and practice of structural equation modeling, The Guilford Press, 4th ed 2016.
12. MacCallum, R.C., Browne, M.W., and Sugawara, H., M., "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," *Psychological Methods*, 1 (2), 130-49, 1996.
13. Hu, L.-t., & Bentler, P. M., Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>, 1998.

Towards On Real Time Early Warning System for Microearthquakes at Cotopaxi Volcano

Román Lara¹[0000-0001-8848-9928], Santiago Altamirano¹[0000-0001-6899-4724], Julio Larco¹[0000-0003-4789-5983], Diego Benítez²[0000-0001-6219-067X], and Noel Pérez²

¹ Universidad de las Fuerzas Armadas – ESPE, WiCOM-Energy Research Group, Centro de Investigación de Redes Ad Hoc, Sangolquí, Ecuador
[ralara, bsaltamirano, jclarco]@espe.edu.ec

² Universidad San Francisco de Quito, Colegio de Ciencias e Ingenierías, Quito, Ecuador
[dbenitez, nperez]@usfq.edu.ec

Abstract. Around the world, volcanic activity has increased and in case of an eruption, the population could be really affected. Ecuador has several active volcanoes, which generates the need to study methods to identify possible eruptions, in order to launch a reliable alert, which allows to safeguard lives and material losses. The aim of this paper is to develop an automatic microseism recognition system, which activates the Early Warning Broadcast System (EWBS). The recognition system is obtained by applying models as k-Nearest Neighbors, Support Vector Machine and Decision Trees by considering frequency features, applied to the data provided by the Instituto Geofísico de la Escuela Politécnica Nacional with the seismic activity presented at Cotopaxi volcano during 2012. For detection stage the system reached a Balanced Error Rate (BER) of 0.01, while in classification stage the system reached a BER of 0.11. Finally, the recognition system is integrated to the EWBS, which presents a delay of 30s for launching an early warning.

Keywords: EWBS, ISDB-T, Recognition system, TDT

1 Introducción

Los sistemas de alerta temprana han demostrado ser una fuente efectiva para salvaguardar vidas frente a eventuales desastres naturales como: derrumbamiento de tierras [1], inundaciones [2], Tsunamis [3][4], aunque aún existe una gran brecha para el caso de terremotos [5, 6] y erupciones volcánicas [7]. Estas últimas representan uno de los peligros naturales permanentemente latentes, que amenazan con el equilibrio de los ecosistemas, por ello los volcanes son permanentemente monitorizados para identificar características relacionadas con la deformación del suelo, flujo de gas, sismicidad y otros factores, y así permitir alertar eficazmente a la sociedad. En este contexto, la sismología es una herramienta importante para reconocer cambios en los volcanes, es así como alrededor del mundo se han definido instituciones encargadas de la vigilancia

volcánica [8-10], en las cuales sus analistas expertos pueden interpretar la firma sismológica para reconocer los diferentes tipos de microsismos que se pueden originar [11] y basados en esta información recomendar la emisión o cambio de alerta [12].

El Ecuador es considerado uno de los países que posee gran cantidad de volcanes activos en el mundo, al encontrarse situado en el Cinturón de Fuego del Pacífico, donde existe una permanente actividad volcánica y sísmica [13]. El Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) es la institución encargada de realizar la monitorización sísmica y volcánica en el Ecuador. Además, el IGEPN es el ente encargado recomendar a la Secretaría de Gestión de Riesgos (SGR) realizar y notificar un cambio de alerta oportuno y eficaz para que la comunidad, con posible afectación, siga las recomendaciones brindadas por el mismo. Actualmente, existen varios volcanes en proceso eruptivo, uno de los volcanes más activos y con mayor peligrosidad es el volcán Cotopaxi, ya que puede afectar alrededor de 300.000 habitantes en caso de una erupción, para lo cual se ha desplegado un sistema de monitorización [14] y mediante un análisis sísmico recomendar un cambio de alerta. Finalmente, emplear el sistema de difusión de alerta temprana (EWBS, del inglés *Early Warning Broadcast System*) basado en el sistema de televisión digital terrestre en concordancia con el estándar de servicio integrado de televisión terrestre (ISDB-T, del inglés *Integrated Services Digital Broadcasting - Terrestrial*) para la emisión de una alerta oportuna, se convierte en una alternativa viable.

En este contexto, el objetivo de este artículo es desarrollar un sistema de alerta temprana frente a una eventual erupción del volcán Cotopaxi, basados en el incremento de sus microsismos. Para cumplir con el objetivo se propone desarrollar un sistema de reconocimiento (detección + clasificación) automático de microsismos empleando técnicas de Machine Learning tradicional como máquinas de vector soporte (SVM, del inglés Support Vector Machine), k vecinos más cercanos (k-NN, del inglés k-nearest neighbors) y árboles de decisión (DT, del inglés Decision Trees). Finalmente, emitir una alerta temprana mediante la activación del sistema EWBS basado en ISDB-T, la definición de los rangos incrementales del número de microsismos que se relacionen con las diferentes alertas definidas en el estado ecuatoriano, está fuera del alcance del presente artículo.

2 Metodología

En la Fig. 1 se observa el sistema propuesto, el cual consta de tres fases generales: detector, clasificador y sistema de alerta temprana. Las bases de datos utilizadas fueron proporcionadas por el IGEPN, el cual ha desplegado 17 estaciones sismológicas en el volcán Cotopaxi. Estas estaciones se encuentran ubicadas en las faldas del volcán, entre las cuales se encuentran 5 estaciones de corto periodo que trabajan en un rango de frecuencia de 1 a 50 Hz y 12 estaciones de banda ancha que trabajan a una frecuencia de 0.1 a 50 Hz [15]. Los datos son pertenecientes a la componente vertical de las estaciones, las cuales trabajan a una frecuencia de muestreo de digitalización de 50 Hz o 100 Hz y recopilados durante varios años. Los microsismos que se presentan más frecuentemente son: de largo período (LP, del inglés *Long Period*), volcano tectónicos (VT,

del inglés *Volcano Tectonics*), temblores (TRE, del inglés *Tremors*) e híbridos (HYB, del inglés *Hybrids*) [16].

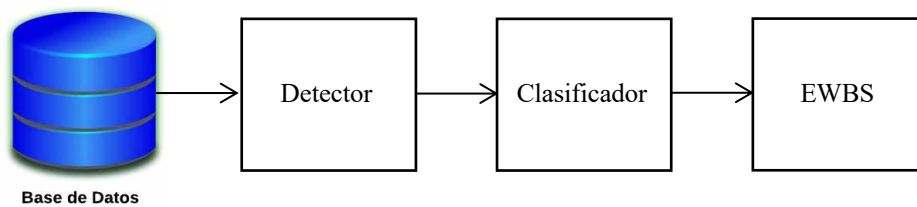


Fig. 1. Diagrama de bloques del sistema de alerta temprana del volcán Cotopaxi.

2.1 Detector

La base de datos empleada para el detector consta de 350 señales de microsismos etiquetados por los expertos del IGEPN, en las cuales posee 282 eventos del tipo LP, 50 eventos del tipo VT, 15 eventos tipo TRE y 3 eventos HYB, estas señales están ordenadas por mes y tipo de microsismo. Las señales entregadas por el IGEPN son señales filtradas en un rango de 1 a 50 Hz, donde previamente han sido eliminados los picos de baja amplitud que se encuentran en frecuencias de 0.1 a 0.3 Hz, debido principalmente a movimientos de placas oceánicas. Las señales recopiladas por la estación VC1 son señales que tienen distintas amplitudes dependiendo de la intensidad de cada microsismo, por lo cual a todas las señales de la base de datos se deben normalizar. Para realizar la normalización se debe remover la tendencia lineal y llevar a cada microsismo a una escala en el rango de [-1,1] centrada en cero, para esto se aplican las siguientes ecuaciones:

$$\hat{r}_l = \frac{r_l - \mu_l}{\sigma_l}, \quad (1)$$

donde \hat{r}_l es la señal normalizada, μ_l y σ_l son la media y la desviación estándar respectivamente y

$$x_i = \frac{\hat{r}_l}{\max|\hat{r}_l|}, \quad (2)$$

donde x_i es la señal estandarizada en un rango entre [-1,1]. En la Fig. 2a se puede observar la comparación de una señal original con otra normalizada.

Como se indicó previamente, la base de datos dispone de eventos sísmicos con una duración de 20 minutos, se propone realizar segmentaciones en ventanas de tiempo. En este trabajo se realiza una segmentación en ventanas de 15s como las realizadas en [16], donde se obtuvieron los mejores resultados en la detección de microsismos con distintas ventanas de tiempo.

4

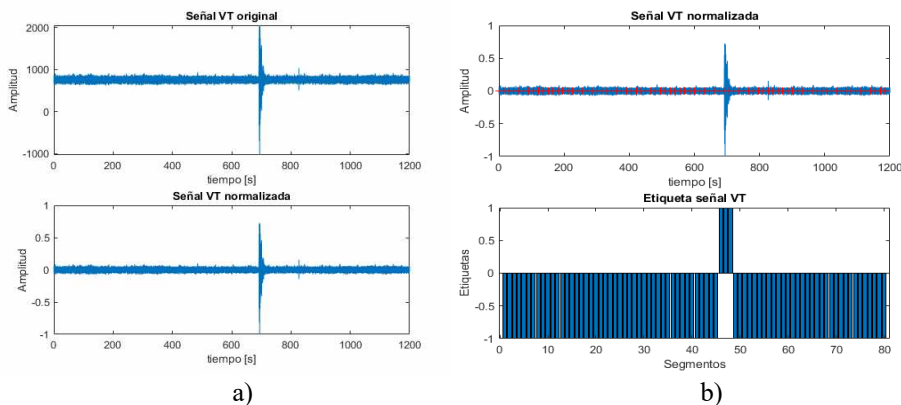


Fig. 2. Ejemplo de un registro que contiene un evento VT. a) registro original versus el normalizado. b) Proceso de segmentación y etiquetado de un registro.

Al realizar una segmentación de 15 segundos se obtienen 80 segmentos de cada microsismo. Estos segmentos son almacenados en un vector s_i y finalmente se agrega cada señal segmentada a una matriz S,

$$s_i = [s_{i,1}^T, s_{i,2}^T, s_{i,3}^T, \dots, s_{i,8}^T]^T, \quad (3)$$

donde $s_{i,j}$ es un segmento de señal de 15s de la señal normalizada x_i y se conforma

$$S = [s_1^T, s_2^T, s_3^T, \dots, s_M^T]^T, \quad (4)$$

donde s_M es un vector de cada señal segmentada y M es el número de señales que se encuentran en la base de datos. Luego se proceden a etiquetar los segmentos que contienen las señales microsísmicas al colocar +1 y los segmentos que no contienen señal con -1. Esta información se encuentra incluida en la base de datos entregada por IGEPN, este proceso se realiza para cada señal y se almacena en un vector definido a continuación:

$$Y_i = [y_{i,1}^T, y_{i,2}^T, y_{i,3}^T, \dots, y_{i,8}^T]^T, \quad (5)$$

donde $y_{i,j}$ es el etiquetamiento de cada segmento de señal $s_{i,j}$,

$$Y = [Y_1^T, Y_2^T, Y_3^T, \dots, Y_M^T]^T, \quad (6)$$

donde Y_M es un vector de cada señal etiquetada y M es el número de señales que se encuentran en la base de datos. En la Fig. 2b se pueden observar los procesos de segmentación y etiquetamiento.

Posteriormente, a cada segmento de señal de 15s se calcula la densidad espectral de potencia (PSD, del inglés *Power Spectral Density*) con el método de Welch. La PSD se calcula con una ventana de 512 puntos y entrega una resolución de 257 puntos, estos puntos de resolución son considerados como características en

frecuencia. Estas características son usadas para el entrenamiento. La PSD de la señal segmentada se almacena en una matriz W , de la siguiente forma:

$$w_i = [w_{i,1}^T, w_{i,2}^T, w_{i,3}^T, \dots, w_{i,8}^T]^T, \quad (7)$$

donde $w_{i,j}$ son las 257 características obtenidas mediante la PSD de cada segmento $s_{i,j}$,

$$W = [w_1^T, w_2^T, w_3^T, \dots, w_M^T]^T, \quad (8)$$

donde W_M es la PSD de cada señal extraída y M es el número de señales que se encuentran en la base de datos.

Finalmente, con la matriz de características W y las etiquetas Y , se procede a realizar la detección de los eventos basados en aprendizaje supervisado, para ello se evalúan varios métodos tradicionales de aprendizaje estadístico como: k-NN, DT y SVM, para cada clasificador se trabaja con un método de selección de características como filtros, embebidos y de envoltura, respectivamente, los cuales son evaluados y optimizados en función de las principales métricas de desempeño del clasificador como son la exactitud (A , del inglés *Accuracy*), precisión (P), sensibilidad (S), especificidad (R , del inglés *recall*) y la tasa de error balanceada (BER, del inglés *Balanced Error Rate*). Con esta estrategia maximizamos la probabilidad de encontrar segmentos que contengan la mayor porción de un microsismo.

2.2 Clasificador

Para esta etapa, la base de datos usada para obtener los modelos de clasificación de microsismos fue recopilada a lo largo del año 2012, 2013, 2014 y 2015. Esta base de datos consta con un total de 1184 microsismos con distintas variaciones de tiempo, esta base está compuesta de 1044 eventos del tipo LP, 101 eventos VT, 27 eventos Regionales, 8 eventos del tipo HYB y 7 eventos de rompimiento de glaciares (ICQ, del inglés *Icequake*) [17]. Las señales han sido filtradas en un rango de 1 a 50 Hz y se procede a realizar la normalización, para ello se debe retirar la tendencia lineal y llevar a cada señal a una escala de $[-1,1]$ centrada en cero, para esto se aplican las ecuaciones (1) y (2). Para el etiquetamiento de las señales se consideran los 1044 microsismos del tipo LP y los 101 microsismos del tipo VT. Solo estos dos tipos de eventos se van a usar para obtener los modelos de entrenamiento para la clasificación de microsismos; para los eventos del tipo LP se coloca una etiqueta de -1 y para los eventos del tipo VT se coloca una etiqueta de +1. En la etapa de extracción de características se calcula la PSD con el método de *Welch* exclusivamente del microsismo, se almacena en un vector como se observa en la ecuación (7). La PSD se calcula con una ventana de 512 puntos y entrega una resolución de 257 puntos, estos puntos de resolución son considerados como características en frecuencia. Estas características son usadas para el entrenamiento del clasificador. La PSD de la señal se almacena en una matriz de acuerdo con la ecuación (8). Finalmente, con la matriz de características W y las etiquetas Y de los microsismos, se procede a realizar la clasificación evaluando k-NN, DT y SVM, para cada clasificador se trabaja con un método de selección de características como filtros, embebidos y de envoltura respectivamente, los cuales son evaluados y optimizados en

6

función de las principales métricas de desempeño del clasificador como son: A, P, S, R y la BER, con esta estrategia obtenemos un clasificador de elevada confiabilidad.

Con la finalidad de mejorar el rendimiento del detector y del clasificador, en cada una de ellas se implementa un sistema de votación, que al trabajar con 3 métodos y realizar la votación, la mayoría gana. En el detector el sistema de votación indica si un segmento es o no parte de un evento, mientras que, para el clasificador, el sistema de votación indica en su mayoría si es un VT o LP.

2.3 Sistema de Alerta de Emergencia

La Universidad de las Fuerzas Armadas – ESPE a través del Ministerio de Telecomunicaciones y en cooperación con el Gobierno de Japón, ha desarrollado un sistema que es capaz de emitir una alerta temprana cuando el sistema de reconocimiento así lo dictamine. A este punto el sistema es capaz de reconocer microsismos LP y VT, para emitir una alerta con fines prácticos, se definieron rangos incrementales de eventos LP y VT y se los relacionó con los tipos de emergencia existentes: verde (entre 0 y 199 eventos), amarilla (entre 200 y 399 eventos) y roja (mayor a 400 eventos).

Para el sistema EWBS, con la finalidad de monitorizar la señal desde el aire y replicar el mensaje enviado por la transmisión EWBS, se implementó un receptor ISDB-T en una arquitectura de radio definida por software de bajo costo, ADALM-PLUTO, de *Analog Devices*. El receptor realiza la función de analizador de las capas física y de transporte, decodificando los códigos de área y el mensaje superpuesto. La alerta se retransmite a través de una red interna, cumpliendo la función de pasarela y enviándola a los dispositivos conectados a la red. Finalmente, se evalúa el tiempo de respuesta del sistema EWBS al emitir un cambio de alerta.

3 Resultados

3.1 Resultados del Detector

Al aplicar las estrategias definidas en la sección anterior se obtuvieron los modelos de los clasificadores con un porcentaje de 60/40 para entrenamiento y prueba, respectivamente, se puede observar en la Tabla 1 los mejores resultados, después de un proceso iterativo de selección de características se identifican 100 características de las 257 como óptimas para los tres clasificadores definidos, con lo que se evita el sobreajuste, además se aplica un sistema de votación que permite maximizar los resultados de las métricas de desempeño alcanzando un BER del 0.01, es decir por que por cada 100 eventos 1 no es detectado.

Tabla 1. Rendimiento del detector basado en clasificadores k-NN, DT, SVM y sistema de votación con 100 características.

Algoritmo	A(%)	P(%)	R(%)	S(%)	BER
k-NN	98	96	98	98	0,02
DT	79	49	98	74	0,13
SVM	94	85	98	95	0,03
Votación	98	98	99	99	0,01

3.2 Resultados del Clasificador

Al disponer en la base de datos de 101 VT se eligieron de manera aleatoria 101 LP con la finalidad de trabajar de forma balanceada, se trabajó con porcentajes de 60/40 para entrenamiento y prueba, respectivamente, es decir 120 microsismos (60 LP +60 VT) para entrenamiento y obtención de los modelos y se probaron con 82 microsismos (41 LP + 41 VT). Además, se realizó una prueba final a los modelos obtenidos, en la que se consideraron 563 microsismos, con 41 VT y 522 LP. Los resultados de los clasificadores después de la fase de selección de características se identifican que son 50 las mejores, identificando a 50 frecuencias que maximizan los valores de las métricas de desempeño, alcanzando una exactitud del 85% y BER del 0.15 al clasificar 82 microsismos, es decir por que por cada 100 eventos 15 son mal clasificados, mientras que, al clasificar 563 microsismos, se mejora en 7% la exactitud y 11 microsismos son mal clasificados. Los resultados se pueden apreciar en la Tabla 2.

Tabla 2. Rendimiento de los clasificadores k-NN, DT, SVM y sistema de votación con 50 características para 82/563 microsismos.

Algoritmo	A(%)	P(%)	R(%)	S(%)	BER
k-NN	82/93	100/54	63/63	100/96	0,18/0,20
DT	83/87	97/32	68/68	97/89	0,17/0,21
SVM	82/87	84/34	78/78	85/88	0,18/0,17
Votación	85/92	100/89	70/84	100/94	0,15/0,11

3.3 Resultados del sistema EWBS

El sistema de reconocimiento automático de microsismos dispone de un contador de eventos, cada vez que este supera un determinado número de eventos el cual está relacionado con una alerta, para nuestro caso práctico se definió como tiempo de observación 1 hora, y por cada 200 eventos reconocidos la alerta cambia de verde a amarilla,

8

al superar los 400 eventos se cambia de alerta de amarilla a roja, y se activa el sistema EWBS, el cual presenta un retardo de 30s en emitir mensajes de superposición en los receptores de televisión con el mensaje de “Amenaza de erupción, orden de evacuación para Sangolquí”. El sistema completo se lo puede apreciar en la Fig. 3.

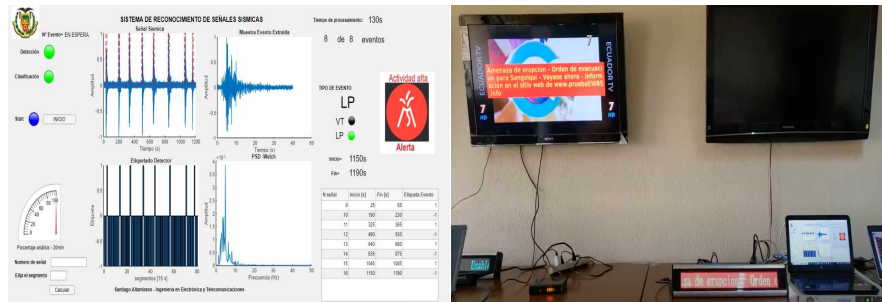


Fig. 3. Sistema EWBS del volcán Cotopaxi.

4 Conclusiones y Trabajos Futuros

Se logró desarrollar un sistema de reconocimiento automático de microsismos del volcán Cotopaxi, donde se obtuvieron modelos de aprendizaje supervisado k-NN, DT y SVM implementados en los sistemas de detección de eventos LP, VT, TR y HYB, y clasificación de eventos LP y VT basados en características en el dominio de la frecuencia, donde se obtuvieron 3 modelos de detección y 3 modelos de clasificación, con estos modelos se realizó un sistema de votación para la detección de microsismos, logrando un BER de 0,01. En el caso del clasificador el sistema alcanza un BER de 0.11.

La activación del sistema EWBS presenta un tiempo de retardo de 30s, el cual es óptimo para este tipo de aplicaciones, que permita en caso de una eventual erupción salvaguardar vidas y recursos de las poblaciones aledañas al volcán Cotopaxi.

Nuestro grupo de investigación está interesado en continuar con la investigación de forma conjunta con el IGEPN y la SGR, para definir los posibles rangos que permitan realizar los cambios de alerta y activar al sistema EWBS oportunamente.

Referencias

1. P. Gaetano, M. Calvello, and L. Piciullo: Monitoring strategies for local landslide early warning systems. *Landslides* 16(2), 213-231 (2019).
2. S. Haritharaj, N. U. Visakh, and A. C. Antony: Early warning systems: An alert to Flood. *Vigyan Varta* 1(1), 30-32 (2020).
3. H. Rahayu, L. Comfort, R. Haigh, et al.: A study of people-centered early warning system in the face of near-field tsunami risk for Indonesian coastal cities. *International journal of disaster resilience in the built environment* 11(2), 241-262 (2020).
4. J. LaBrecque, et al.: Global navigation satellite system enhancement for tsunami early warning systems. *Global Assessment Report on Disaster Risk Reduction* (2019).

5. Zhang, Meng, et al.: Brief communication: Effective earthquake early warning systems: appropriate messaging and public awareness roles. *Natural Hazards and Earth System Sciences* 21(10), 3243-3250 (2021).
6. W. Adi, et al.: Earthquake early warning system using ncheck and hard-shared orthogonal multitarget regression on deep learning. *IEEE Geoscience and Remote Sensing Letters* 19, 1-5 (2021).
7. A. Shin, et al.: MOWLAS: NIED observation network for earthquake, tsunami and volcano. *Earth, Planets and Space* 72(1), 1-31 (2020).
8. R. Sparks, J. Biggs, and J. Neuberg: Monitoring volcanoes. *Science*, 1310-1311 (2021).
9. J. Lowenstern, et al.: Guidelines for volcano-observatory operations during crises: recommendations from the 2019 volcano observatory best practices meeting. *Journal of Applied Volcanology* 11(1), 1-24 (2022).
10. J. Pallister, et al.: Volcano observatory best practices workshops-a summary of findings and best-practice recommendations. *Journal of Applied Volcanology* 8(1), 1-33 (2019).
11. R. Lara-Cueva, D. Benítez, E. Carrera, M. Ruíz y J. Rojo: Automatic Recognition of Long Period Events From Volcano Tectonic Earthquakes at Cotopaxi Volcano. *IEEE Transactions on Geoscience and Remote Sensing*, 54(9), 5247-5257 (2016)
12. O. Corominas y J. Martí: Estudio comparativos de los planes de actuación frente al riesgo volcánico (Chile, Costa Rica, El Salvador, Ecuador, España, México y Nicaragua), *Revista Geológica de América Central*, 2011.
13. IGEPN, «Volcanes más activos del Ecuador,» 2020. [En línea]. Available: <https://www.igepn.edu.ec/>.
14. H. Ortiz Herazo, «Estudio de los Efectos de Sitio para la Construcción de un Índice de Actividad Sísmica en el volcán Cotopaxi,» Escuela Politécnica Nacional, 2013.
15. IGEPN, «Cotopaxi,» 2017. [En línea]. Available: <https://www.igepn.edu.ec>.
16. R. Lara, D. Benítez, V. Carrera, M. Ruiz y R. Rojo Álvarez: Feature selection of seismic waveforms for long period event detection at Cotopaxi Volcan. *Journal of Volcanology and Geothermal Research*, 316, 34-49 (2016).
17. N. Pérez, et al: ESeismic: Towards an Ecuadorian volcano seismic repository. *Journal of Volcanology and Geothermal Research* 396, 106855 (2020).

Encryption of messages on the transport stream of Digital Terrestrial Television

Evelina Silva¹, Nelson Benavides¹ and Gonzalo Olmedo¹

¹ Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador
WiCOM-Energy Research Group
[mesilva, gfolmedo, nbbenavides1]@espe.edu.ec

Abstract. Within emergency support systems, digital terrestrial television is of great help. It can send additional information to support personnel using its transport stream through file communication protocols or text, for which special receivers are designed for this purpose. This article proposes a message encryption algorithm in the spaces not used by the PSI/SI tables of the transport stream of the MPEG2-TS standard in such a way as not to affect the content of audio, video, data, or original table configuration.

Keywords: Encryption, TS, TDT, ISDB-T.

1 Introduction

Traditionally, the digital terrestrial television signal transmission system under the ISDB-T standard is carried out using a Transport Stream (TS) [1]. The TS comprises packets with a size of 188 bytes [2]. The modulation and Reed Solomon error correction stage are used; these packets have a dimension of 204 bytes, taking BTS's name (Broadcast Transport Stream) [3]. Each package of a TS or BTS presents specific information about each PSI/SI table belonging to the ISDB-T standard [4].

To transmit the audio, video, and data signals in the TS or BTS file, there is a fixed bandwidth of 29.958 MHz; since the entire bandwidth is not used with the video, audio, and data information, it is you must pad the TS or BTS with null packets to make it reach the set bandwidth [5]. For this reason, a TS or BTS file will have many null packets within its structure; in the same way, within the PAT and PMT tables, there will be padding with a value of 0xFF in hexadecimal to complete the 188 bytes in TS and 204 bytes in BTS [6].

Within the ISDB-T standard, the EWBS (Emergency Broadcasting System Emergency) can be transmitted, which is still in development and evaluation its implementation in Ecuador [7].

One way to support this system is to encrypt any information within the packets containing the PAT and PMT tables, where the null bytes are modified so that better emergency information can be transmitted within the TS or BTS packets. In this sense, the objective of this article is to encrypt specific information, such as video, images, audio, text, etc., in the null packets of the PAT and PMT tables inside a TS or BTS file containing that, which can help to receive more emergency information. Therefore, the occurring emergency can be visualized more clearly by deciphering the report.

2

This article is constituted as follows: in section 2, there is the methodology used for the encryption and decryption of information. Section 3 presents the results obtained, and section 4 the conclusions.

2 Methodology

The present work is intended to encrypt digital content for video, audio, text, etc.; it is divided into two parts. First, a transmission algorithm must be presented, identifying the PAT and PMT tables within the TS or BTS file and replacing their null files with the desired information content. Second is the reception algorithm, where the digital content is reconstructed.

2.1 Transmission algorithm

For the transmission algorithm, a file with a TS extension is selected in read mode to be able to access basic information about the element, such as its name and the size of bytes that, when divided by 188, will have the number of packets that the file has in its structure. Next, each package that makes up the TS file is analyzed, and the PAT and PMT tables are located. With these files located, we enter these tables using their PID and insert the content of a digital file.

To carry out this activity, the structure of the PSI/SI Tables should be considered, as shown in Figure 1.

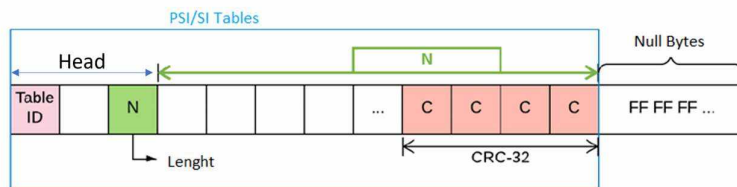


Fig. 1. Structure of the PSI/SI tables.

The PSI/SI tables have their identification that differentiates what content is found in each packet. The length of the information that each package carries is displayed, and finally, it is filled in null bytes with a value of 0xFF in hexadecimal. This is where the encrypted information will be inserted into each PAT and PMT table. Next, the structure of the PAT and PMT tables will be presented.

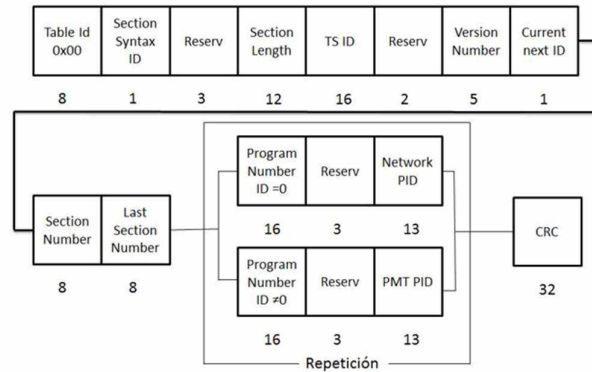


Fig. 2. Structure of the PAT table.

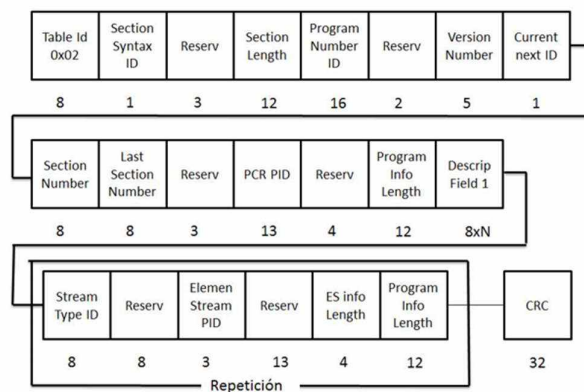


Fig. 3. Structure of the PMT table.

This information presented in Figures 2 and 3 is called the payload, with a total size of 184 bytes, and is inside the green lines in Figure 1. Before that is the packet header with a length of 4 bytes. When the packet information is complete in the payload, the bytes are filled with a value of 0xFF in hexadecimal, the null bytes, with a total of 184 bytes. This is where the information of the digital content is inserted.

To be able to insert a digital file, it must meet specific size characteristics delimited mainly by the number of padding bytes available in the PAT and PMT tables. To ensure reception, it is recommended to be sent at least three times within the TS. This is because tuning to a DTT channel is not always executed from the first moment, which generates a loss of packets and information.

To calculate the size of the digital file as a function of the duration time of the TS file and the padding bytes available between the PMT and PAT tables, it is necessary to use the following equation:

4

$$\eta = \frac{x t_{\text{TS}}}{3 t_t} \quad (1)$$

where:

- η is the size of the file to be encrypted expressed in bytes.
- t_{TS} , is the duration time of the TS
- x is the number of available padding bytes summed across the PAT and PMT tables.
- t_t , corresponds to 100 ms, defined in the standard as the time in which the PAT and PMT tables are transmitted [8].

Then, the byte string of the digital file is opened, and it proceeds to insert the null bytes into the PAT and PMT tables. To do this, a 5-byte digital content header is inserted where the first 3 bytes indicate the packet number, and the 2 bytes are divided into 14 bits for packet size and two flag bits set out in Table 1. The function of the header is to allow the receiver to execute the decryption process until all the packets have been recovered, that is until all the digital content is reconstructed.

Subsequently, the byte string is entered after inserting the header into the null bytes of the PAT and PMT tables, as shown in Figure 4. When finished, a new TS file is created with the encrypted information, ready to be transmitted.

Table 1. Indicator.

<i>Valor</i>	<i>Indicator</i>
10	Start
00	Transmission
01	End

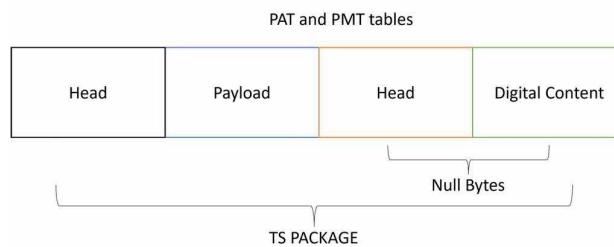


Fig. 4. Structure of the PAT and PMT tables with the encrypted message.

2.2 Reception algorithm

The reception algorithm is intended to reconstruct the encrypted digital content in the null bytes of the PAT and PMT tables. With the help of the header created in the transmission algorithm, the content can be written in vector form later in a new digital file.

Initially, the TS file is opened with information about the number of bytes and packets. Then, the PIDs of the PAT table and the PMT table are searched, and the digital content package is recovered; for this purpose, a function called ANALISIS is created that allows analyzing the header of the digital file with a 5-byte vector that indicates if said packet contains end of digital content information.

To reconstruct the digital content, there are two cases: the first, the reception runs from the first moment of transmission, and the second, where the reception runs from an unknown moment, so the final packet is found before the first.

The received file is opened and stored in a variable named TRAMA in the PAT and PMT Tables to solve these inconveniences. Using the ANALISIS function, it is verified when the beginning and end of the digital content begin again, resulting in the correct order to decrypt the file correctly. Finally, the decrypted file is created to view and compare it with the original file.

3 Results

Once the TS file has been generated with the encrypted data, said signal is transmitted and received, for this the DEKTEC DTA-115 modulating card and the Stream Xpress software are used to carry out the transmission, and the RANGER NEO 2 equipment that will receive the TS file through an antenna and proceeds to capture the signal and save it for analysis. Once the call is charged, the TS file is copied for research.

To begin with, we analyze the encrypted TS file and verify the information to be transmitted using a TS and BTS analyzer. This configuration is displayed in Figure 5, where the encryption of an image is presented in the PAT table, and Figure 6, where it shows the encryption of the image in the PMT table.

6

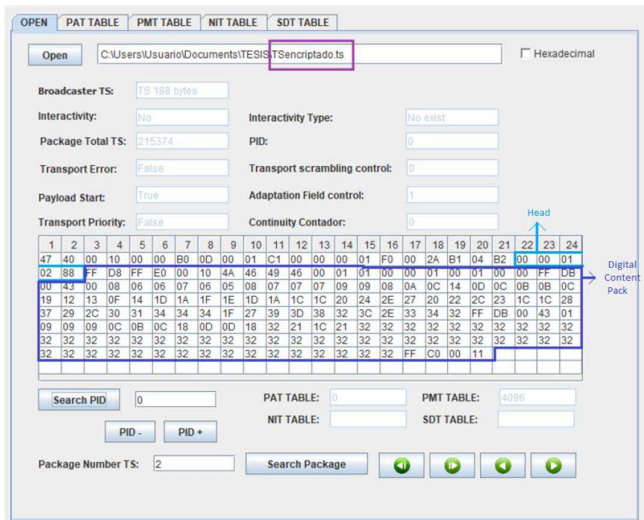


Fig. 5. Structure of the PAT and PMT tables with the encrypted message.

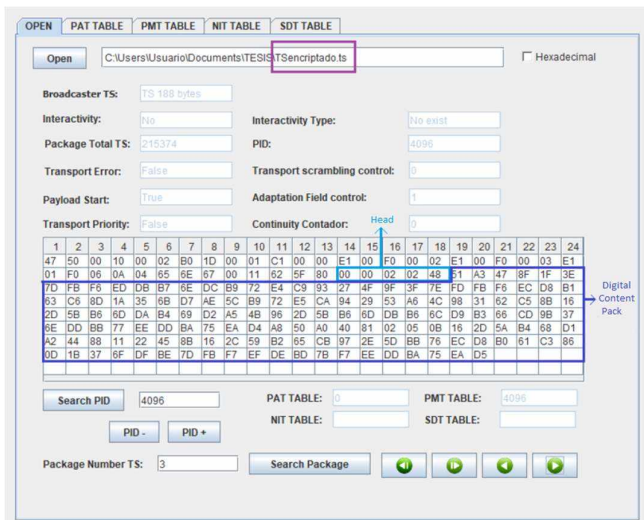


Fig. 6. Structure of the PAT and PMT tables with the encrypted message.

To search for each table, the value of 0 corresponding to the PAT table and the value of 4096 corresponding to the PMT table are inserted in the PID search engine, as shown in Figures 4 and 5, respectively, in turn, shown in the boxes Light blue is the header of the encrypted data and in the purple box the bytes of the digital content.

When the new TS file from the NEO RANGER 2 equipment is obtained, its content is displayed with a TS and BTS analyzer where the order differs from the transmission. That is why the function ANALYSIS and the MATLAB software proceed to order employing a variable that will store the encrypted file to be able to decrypt it without inconvenience. Once this is done, the content of the packets in the PAT and PMT tables is displayed as indicated in Figure 7 and Figure 8.

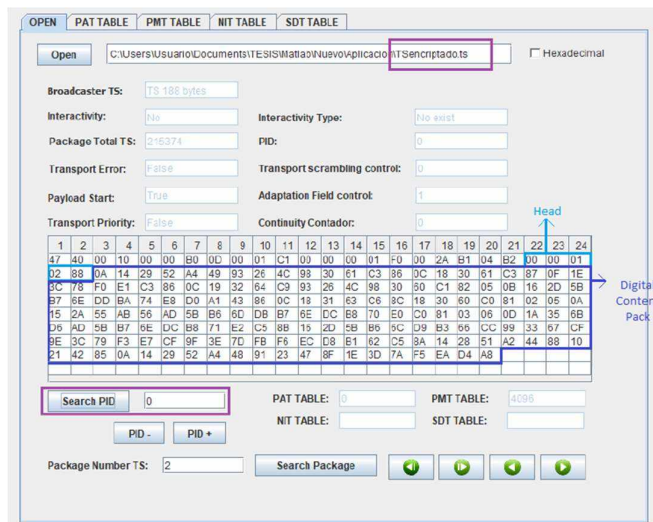


Fig. 7. PAT table with vector content pack.

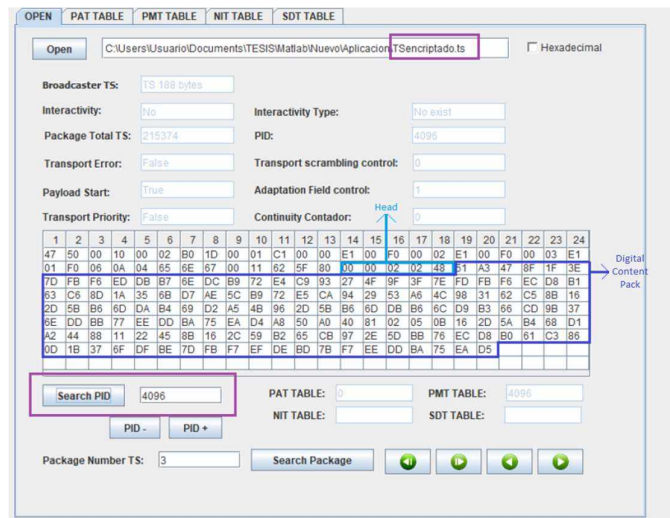


Fig. 8. PMT table with vector content pack.

8

What is displayed in Figures 7 and 8 is the result of ordering the received signal with a vector in an undetermined order; that is why the image is transmitted three times within the PAT and PMT tables to search for the start of the digital data and to be able to order it in the vector so that the transmitted image can be displayed.

In the case of sending the image, the original audio and video signal is not altered, having a good performance at the reception time. Likewise, the encrypted image can be viewed in a file generated from a picture in MATLAB, as shown in the image Figure 9.

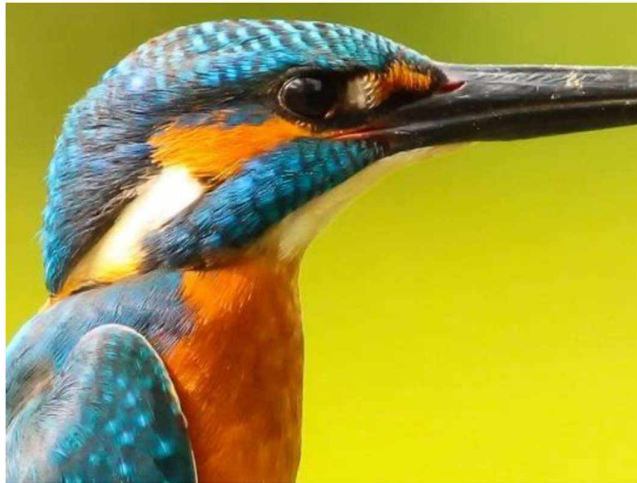


Fig. 9. Decrypted image.

Figure 9 is the result of transmitting an encrypted and decrypted image in the null bytes of the PAT and PMT Tables. In the same way, any digital information that is necessary to information can be sent, as is the case of early emergency alerts that, by encrypting the data, can be sent more quickly to the population.

4 Conclusions

This document has demonstrated an encryption and decryption message in a transport stream using the null bytes of the PAT and PMT tables to transmit digital content, be it text, video, image, audio, etc. The communication system was established through the Dektec transmitter and the Neo Ranger 2 equipment as a receiver, which will capture the encrypted signal. To encrypt the digital data and decrypt MATLAB, when creating functions, the TS file is modified, and this message is extracted using the PID from the PAT and PMT tables using the header information that was inserted in the null bytes of digital components.

In addition, a new way of transmitting information for emergencies and other types of data that may be required in a communication system has been carried out where it may be imperceptible to the user, and a suitable receiver can generate this information to alert the population in the case of emergency early warning.

The current work is focused on sending any information using null bytes to take advantage of the channel bandwidth.

References

1. R. Sotelo, "Sistema de transmisión ISDB-T," *Memoria Investigaciones en Ingeniería*, no. 9, pp. 67–77, 2011.
2. D. Villamarín, G. Olmedo, R. Lara, and M. A. Illescas, "Generación de Transport Stream con Audio, Vídeo y Datos de Interactividad para el Sistema de Televisión Digital Terrestre ISDB-Tb," *MASKAY*, vol. 2, no. 1, pp. 49–55, 2012.
3. J. L. Mendoza Torres, "Diseño e implementación de un banco de pruebas de laboratorio aplicado a la televisión digital utilizando SLINGBOX, VLC y análisis de la capa física del estándar ISDBT con radios definidas por software," 2022.
4. N. Pisciotto, "Sistema ISDB-Tb (primera parte)," *Publicaciones de la Universidad Blas Pascal. Serie Materiales de Investigación*, vol. 3, 2010.
5. J. A. Peñaherrera Obregón, "Implementación de un laboratorio de pruebas de TV digital terrestre ISDB-Tb, para la emisión de aplicaciones interactivas para la escuela de ingeniería electrónica y telecomunicaciones de la UNACH," 2014.
6. J. R. Camalle Vásquez and G. F. Olmedo Cifuentes, "Desarrollo de un algoritmo de compresión del transport stream enfocado en redes de contribución bajo el estándar ISDB-T".
7. D. P. Villacrés Jiménez, "Artículo Científico-Implementación de un sistema piloto de transmisión de alerta de emergencia sobre la televisión digital terrestre en el ECUADOR," 2013.
8. P. E. Enríques Chicaiza and E. D. Rivadeneira Obregón, "Análisis comparativo para el servicio de televisión digital terrestre en Quito," 2018.

Incorporando *eye tracking* no sistema de visualização de ondas neurais baseado em EEG

Matheus Cavalcanti¹[0000-0002-0094-7355], Felipe Melo¹[0000-0003-3930-5760], Thiago Silva¹[0000-0003-1689-1736], Matheus Falcão¹[0000-0001-6441-3144] and Valdecir Becker¹[0000-0002-0639-3875]

¹ Laboratory of Interaction and Media, Informatics Center,
Federal University of Paraíba, João Pessoa - PB, Brazil
contato@lim.ci.ufpb.br

Abstract. Este artigo descreve a incorporação do *eye tracking* no sistema de visualização e análise das ondas neurais, baseado em eletroencefalografia (EEG), durante a fruição de conteúdos audiovisuais. O sistema de visualização foi desenvolvido em Python, utilizando o *headset* Emotiv Insight. Durante os testes, percebeu-se a necessidade de identificar se as reações mapeadas pelo EEG estavam de fato relacionadas à fruição do conteúdo ou se tinham origem em elementos externos à tela, com o indivíduo desviando o olhar e, conseqüentemente, a atenção. Baseado na metodologia do *Design Science Research*, o *eye tracking* foi incorporado à arquitetura do sistema. Para a validação, foram realizados testes com 10 usuários, utilizando um *trailer* de filme. Analisando os dados gerados, foi possível identificar a correlação entre as informações apresentadas pelo EEG e o olhar dos indivíduos. Dessa forma, é possível elevar a certeza sobre a origem das emoções durante a fruição de conteúdos audiovisuais.

Keywords: Eye Tracking, EEG, Fruição Audiovisual.

1 Introdução

A tecnologia de *eye tracking* é bem conhecida e utilizada em diversas áreas do conhecimento, tais como pesquisas de usabilidade, marketing, experimentos da psicologia cognitiva, ou como entretenimento, em jogos. Seu funcionamento se baseia na detecção do movimento dos olhos de um usuário durante a interação. Isso pode ser feito através de diferentes ferramentas: *webcams*, câmeras infravermelhas, displays montados na cabeça. Os dados coletados fornecem informações relevantes sobre os usuários, como a sua atenção em um determinado objeto na tela, por exemplo.

A tecnologia do *eye tracking* é flexível e pode ser utilizada em conjunto com outras ferramentas, como leitores e visualizadores de ondas neurais. Em etapas anteriores desta pesquisa, foi desenvolvido um sistema de visualização de ondas neurais baseado em eletroencefalografia (EEG) (9, 10). O objetivo é mapear emoções durante o processo de fruição de conteúdos audiovisuais a partir de alterações nos padrões das ondas neurais. No entanto, um problema que se apresentou durante os testes estava relacionado ao foco das emoções, que em alguns casos tinham origem externa à tela da TV. Como exemplo específico, em um determinado momento o telefone celular de um dos participantes tocou, gerando uma atenção e um estado de estresse, identificados

pela medição das ondas neurais. No entanto, para a finalidade do sistema, só emoções oriundas da fruição audiovisual são pertinentes.

Visando sanar essa problemática, foi incorporado ao sistema o *eye tracking* por meio de *webcam*, que permite identificar quais elementos de uma cena o usuário está olhando ou se sua atenção foi desviada para outro assunto fora da tela. Para os testes, foi apresentado ao usuário um *trailer* de filme, enquanto era feita a medição de suas ondas cerebrais através do dispositivo de EEG Emotiv Insight. Dessa forma, o objetivo deste artigo é demonstrar como o uso do *eye tracking* auxiliou na identificação dos pontos específicos da tela que chamaram a atenção do usuário, indicando o que possivelmente causou os resultados observados por meio da visualização das ondas do EEG.

O resultado demonstra uma correlação clara entre os momentos de atenção gerados pelo filme e o olhar dos indivíduos. Ao analisar a atenção identificada pela leitura das ondas neurais e compará-la com o direcionamento do olhar, é possível inferir claramente quais elementos filmicos, e como foram organizados na tela, despertou a curiosidade do espectador. Dessa forma, é possível ter um índice maior de certeza de que a atenção, e conseqüentemente, as emoções sentidas pelos indivíduos, foram de fato gerados pelo filme e não por elementos externos à tela.

2 O método *Design Science Research*

No presente artigo, foi utilizado um software de *eye tracking* para *webcam*, com o objetivo de auxiliar na identificação de pontos de interesse do usuário nas cenas de um *trailer* exibido na tela. O desenvolvimento da pesquisa está baseado na Design Science Research (DSR). Esse processo metodológico legitima o desenvolvimento de artefatos orientados à resolução de problemas como uma maneira importante de produzir conhecimento científico e tecnológico [11,12,13]. A DSR foi considerada relevante para este estudo após a identificação de lacunas na produção técnica e científica no campo dos sistemas audiovisuais. Para a DSR, um elemento fundamental é a compreensão de ambientes externos e internos que envolvem o artefato a ser criado. Verificou-se que, nos últimos anos, não há menção, nem na IHC, nem nos estudos de mídia, à criação de modelos, métodos ou estruturas capazes de apoiar a visualização gráfica das ondas neurais durante a fruição de conteúdos de sistemas audiovisuais com mapeamento do olhar [14]. Os estágios desse método consistem em seis atividades: 1: Identificação e motivação de problemas; 2: Definição dos objetivos de uma solução; 3: Design e desenvolvimento; 4: Demonstração; 5: Avaliação; e 6. Comunicação.

Neste artigo é descrito o uso da ferramenta GazeRecorder, um software online gratuito para experimentos de *eye tracking* baseado na câmera do computador. Para as etapas de demonstração e validação, um experimento foi montado com um *trailer* de filme. Foi feita uma etapa inicial de calibração com o usuário, que consiste em seguir, com a visão, um ponto vermelho que se desloca sobre a tela. Após a calibração, o *trailer* foi exibido na tela para o usuário, que teve sua visão mapeada. Dessa forma, a ferramenta foi capaz de identificar regiões do vídeo que o usuário teve maior concentração. Finalmente, o software gera um mapa de calor sobre as cenas do filme, dispondo os resultados do experimento para análise posterior. O software utilizado possui acurácia de 1,05°, precisão de 0,129° e frequência de amostragem de 30Hz [1].

3 Eye Tracking e trabalhos relacionados

Na literatura, a combinação do EEG com ferramentas de *eye tracking* para melhorar técnicas de reconhecimento de emoções pode ser observada em Lu et al., mostrando-se positiva, com acurácia de 87,59% usando uma integral fuzzy para realizar a fusão dos dados [2]. Em Lopez-Gil et al., vemos outra aplicação para aprimorar o reconhecimento de emoções baseado em EEG, *eye tracking* e outros sinais biométricos [3]. No contexto dos sistemas de recomendação de conteúdo, Xu et al. utiliza dados referentes à atenção do usuário a itens em uma plataforma online de documentos, imagens e vídeos, obtidos por meio de *eye tracking*, para recomendar novos conteúdos de forma personalizada [4]. O *eye tracking* também pode ser visto em pesquisas de usabilidade [5] e em neuromarketing [6], como uma ferramenta para prever o padrão de uso/consumo do indivíduo.

Para uma melhor compreensão de como funciona a tecnologia de *eye tracking* e como ela pode fornecer dados relevantes para a análise do interesse, gosto ou estado emocional do usuário, é importante entender alguns termos que ajudarão a definir quais tipos de informações serão observadas quando realizar um experimento de *eye tracking* (FIG. 1).

O primeiro termo em questão é o “gaze point”, que se refere aos pontos específicos da imagem que o usuário olhou. Um conjunto de *gaze points* agrupados em um determinado espaço em um certo período é chamado de “fixação” [7]. Essa região é de suma importância para nosso estudo, pois é durante sua ocorrência que ocorrem os principais processos cognitivos do usuário, como compreensão e memória.

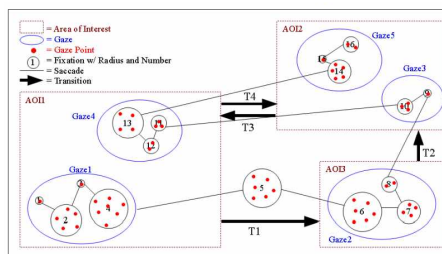


Fig. 1. Representação dos termos relacionados à compreensão do *eye tracking* [7].



Fig. 2. Mapa de calor obtido por meio de *eye tracking* [8].

O movimento rápido realizado entre as áreas de “fixação” é chamado de “saccade”. Durante sua ocorrência, que corresponde entre 30 e 80 ms, a informação visual é suprimida. Um conjunto de “fixações” pode ser agrupado por proximidade em uma “gaze”, que por sua vez se organiza em “áreas de interesse” (AOI) [7]. O tempo gasto em cada AOI (*dwell time*) pode ser um fator que definirá o interesse do usuário em determinado estímulo apresentado na tela, e um tempo maior gasto pode significar um maior nível de interesse. Uma forma de observar esse fenômeno, que foi utilizada neste trabalho, é por meio de mapas de calor (*heatmaps*), que apresentarão na tela os pontos que chamaram a atenção do usuário [8], conforme mostra a Fig. 2.

Utilizando as informações indicadas acima, associadas aos dados obtidos por meio do EEG, é possível gerar um diagnóstico quanto à percepção do usuário sobre os estímulos visuais. Isso permite compreender melhor as preferências e os processos emocionais e cognitivos que ocorrem naturalmente à medida que o conteúdo em uma tela é consumido.

4 Resultados e discussão

Para a validação do artefato, foi realizado um experimento envolvendo 10 indivíduos, todos do sexo masculino, em sua maioria jovens e pessoas de meia idade, integrantes da comunidade acadêmica do Centro de Informática da Universidade Federal da Paraíba. Os usuários foram expostos ao *trailer* do filme Telefone Preto, que tem como principal característica o terror.

Durante a fruição é possível perceber que quando apareciam legendas na tela, os usuários focaram sua visão nesses elementos textuais (Fig. 3), apresentando uma transição rápida dos elementos cinematográficos para as legendas (*saccades*).

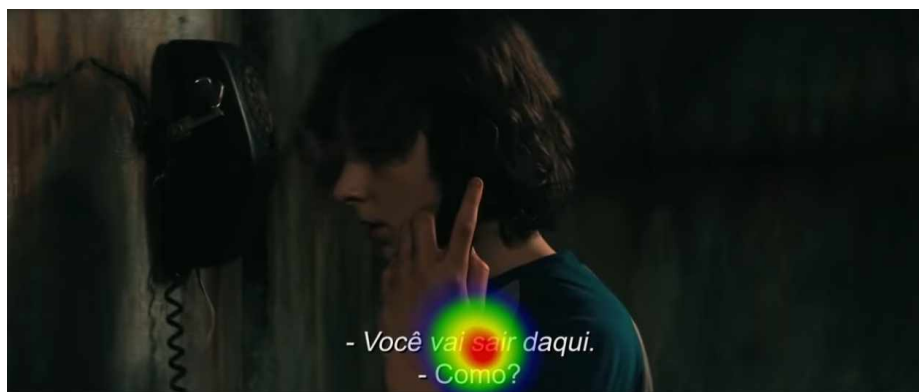


Fig. 3. - Foco ocular na legenda [Autoria própria]

Quando não há um único elemento em evidência, e sim diversos objetos compondo a cena, é perceptível uma dispersão maior dos pontos de calor no *eye tracking* (Fig. 4). Isto pode indicar que o usuário esteja buscando algum elemento crucial da cena, ou aproveitando o momento para compreender o ambiente em que a narrativa ocorre.



Fig. 4. Dispersão dos pontos de calor [Autoria própria].

Fazendo uma correlação com o EEG, em momentos de *jumpscare* (saltos inesperados de algo que possa causar susto no telespectador), os indivíduos tiveram um elevado índice de ondas Low beta (onda que pode ser interpretada por medo, tensão, sustos de forma negativa). Comparando o *eye tracking* com esses momentos, é possível inferir que o usuário levou um susto por causa do elemento que estava em cena, pois houve um alto foco ocular neste personagem em evidência (Fig. 5).



Fig. 5. Foco ocular em personagem em evidência [Autoria própria]

Outra correlação que podemos fazer entre o resultado do *eye tracking* com o EEG é na cena mostrada na Fig. 6. Aos 29 segundos do *trailer*, podemos notar a presença de uma área de interesse disposta sobre o elemento central da cena. A figura do personagem aparece de maneira obscura e misteriosa, aumentando os níveis de tensão do espectador. Isso pode ser observado no gráfico da Fig. 7, que mostra uma grande atividade das ondas Low Beta e High Beta nesse mesmo momento. Podemos então deduzir que o participante estava ansioso ou com medo da cena, com foco no personagem e em suas ações.



Fig. 6. Eye Tracking na cena aos 29 segundos [Autoria própria]

Por meio desse tipo de análise, pudemos observar a sincronia entre a presença de personagens que se mostraram, por meio do EEG, causar medo e tensão nos participantes e a presença de *fixations* sobre sua figura no *eye tracking*. Dessa forma, confirmamos que a causa dos resultados verificados nos gráficos de EEG se deram por elementos específicos em cena.

Sendo assim, a associação entre o EEG e o *eye tracking* nos permite inferir informações e características sobre os usuários que não seriam detectadas por meio do uso de apenas uma das soluções. Ou seja, a distinção entre os elementos cenográficos que provocaram as reações nos espectadores se mostra importante para uma futura classificação dos usuários e das cenas, com base em seus atributos.

5 Conclusão

Este artigo demonstrou o uso do *eye tracking* como ferramenta auxiliar em um projeto de pesquisa utilizando dispositivo de EEG para identificação de emoções. É possível afirmar que os resultados do *eye tracking* foram cruciais na análise das ondas cerebrais, pois permitiu associar os elementos cinematográficos que chamaram a atenção do usuário, possibilitando um paralelo entre suas ondas cerebrais com o mapa de calor gerado. Com isso, também possibilitou compreender como o usuário responde a determinados estímulos visuais. De modo geral, é possível afirmar que o *eye tracking* atendeu ao objetivo proposto e se demonstrou uma excelente ferramenta auxiliar.

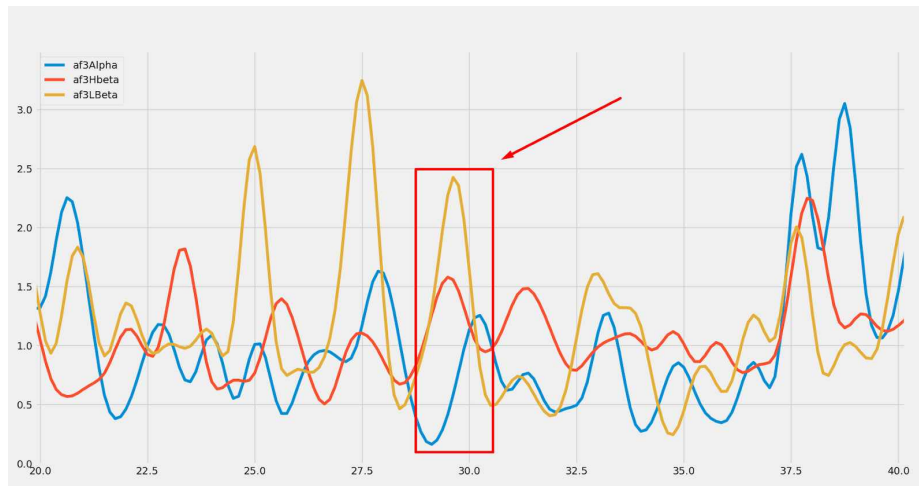


Fig. 7. Ondas Alpha, High Beta e Low Beta, com pico de atividade em evidência aos 29 segundos [Autoria própria]

Como limitações da pesquisa, podemos ressaltar a falta de um ambiente mais confortável para o usuário, e com menos distrações, uma vez que os testes foram feitos em laboratório de pesquisa, sem a emulação de um ambiente de sala de estar ou de cinema. Outro fator limitante é a utilização da webcam do notebook, sendo esta considerada inferior a outros modelos disponíveis no mercado, e que apresentou pequenos problemas de calibração. Estes problemas de calibração foram suspeitados por momentos em que o foco ocular no mapa de calor estava um pouco mais à esquerda ou à direita de algum elemento em evidência na cena. Além disso, os testes foram realizados com apenas 10 indivíduos, com características similares em termos de idade e formação.

Portanto, como melhoria para trabalhos futuros, podemos elencar uma câmera superior para os testes, como por exemplo, baseada em infravermelho. Outro ponto de melhoria é um ambiente mais confortável e que proporcione uma menor distração do usuário, para que não haja perda de informação durante o processo de fruição. Além disso, será preciso repetir os testes com maior quantidade de indivíduos, visando identificar padrões de leitura neural que permitam analisar em profundidade as emoções mapeadas.

Referências

1. SIMPLY USER, User Experience Lab. The comparison of accuracy and precision of eye tracking: GazeFlow vs. SMI RED 250, 2013.
2. LU, et. al. Combining Eye Movements and EEG to Enhance Emotion Recognition, Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015), 2015.
3. LÓPEZ-GIL, et. al. Method for Improving EEG Based Emotion Recognition by Combining It with Synchronized Biometric and Eye Tracking Technologies in a Non-invasive and Low Cost Way. Front. Comput. Neurosci, 2016.
4. XU, et. al. Personalized Online Document, Image and Video Recommendation via Commodity Eye-tracking, RecSys'08, 2008.

5. POOLE, Alex; BALL, Linden. Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects, 2010.
6. SANTOS, et. al. Eye Tracking in Neuromarketing: A Research Agenda for Marketing Studies, *International Journal of Psychological Studies*, Vol. 7, No. 1, 2015.
7. BLASCHECK, et. al. State-of-the-Art of Visualization for Eye Tracking Data, *Eurographics Conference on Visualization (EuroVis)*, 2014.
8. FARNSWORTH, Bryn. 10 Most Used Eye Tracking Metrics and Terms, *iMotions*, 2020.
9. DA SILVA, Thiago Henrique Coelho Tavares; CAVALCANTI, Matheus Dantas; BECKER, Valdecir. Desenvolvimento de um sistema para visualização gráfica de ondas neurais durante consumo de conteúdos midiáticos. IV Jornada Internacional GEMInIS (JIG 2021)-[https://doity.com.br/anais/jig2021/\(03 2022\)](https://doity.com.br/anais/jig2021/(03%2022)). <https://www.doity.com.br/anais/jig2021/trabalho/227774>, 2022.
10. BECKER, Valdecir et al. A System for Graphical Visualization of Brainwaves to Analyse Media Content Consumption. In: *International Conference on Human-Computer Interaction*. Springer, Cham, 2022. p. 318-328.
11. DRESCH, Aline; LACERDA, Daniel Pacheco; JÚNIOR, José Antonio Valle Antunes. Design science research: método de pesquisa para avanço da ciência e tecnologia. *Bookman Editora*, 2020.
12. JÄRVINEN, P. Action research is similar to design science. *Quality & Quantity*, 41, 1 (2007), 37–54.
13. HEVNER, A.R.; MARCH, S.T.; and PARK, J. Design research in information systems research. *MIS Quarterly*, 28, 1 (2004), 75–105.
14. TOSCANO, R. M. et al. HCI Methods and Practices for Audiovisual Systems and Their Potential Contribution to Universal Design for Learning: A Systematic Literature Review. In: *International Conference on Human-Computer Interaction*. Springer, Cham, 2019. p. 526-541.

Improvement of network protocol and analysis of security using aspect of cryptography

Nisarg Patel¹[0000-0001-9083-3660], Viral Parekh²[0000-0002-5774-4049] and
Kaushal Jani³[0000-0003-2284-1110]

¹ Faculty of Technology and Engineering, C. U. Shah University, Wadhwan city, Gujarat, India.

nisargce31@gmail.com

² Department of Computer Engineering, Sal Engineering and Technical Institute, Ahmedabad, Gujarat, India. viral.ccet@gmail.com

³ Department of Computer Sci. Engineering, Indus University, Ahmedabad, Gujarat, India.

drkmjani@gmail.com

Abstract

Network protocol is one of the most discussed topics and used of quantum cryptography is adding value to the network performance. In this research paper we are going to make a novel method to create an implementation of quantum cryptography for key distribution in 802.11 networks. They're convenient due to their portability and high-speed information exchange in homes, offices and enterprises. The advancement in modern wireless technology has led users to switch to a wireless network as compared to a LAN based network connected using Ethernet cables. Unfortunately, they often aren't secure and security issues have become a great concern. If someone is connected to a Wi-Fi network, and sends information through websites or mobile apps, it is easy for someone else to access this information. Based on the laws of physics, quantum cryptography allows exchange of cryptographic key between two remote parties with unconditional security. The foundation of quantum cryptography lies in the Heisenberg uncertainty principle, which states that certain pairs of physical properties are related in such a way that measuring one property prevents the observer from simultaneously knowing the value of the other. In our current research paper, we are using quantum cryptography with AOMDV routing protocol and measure the network performance so using quantum cryptography we will have both secured and improved performance in wireless adhoc network.

Keywords: Cryptography, AOMDV, Quantum Channel, Key Distribution, Quantum Cryptography, Cryptography Techniques

1. INTRODUCTION

Everyone wants their essential information to be readily available, portable, and reachable from practically everywhere they go today, and using wireless networks makes this feasible. As their name implies, wireless networks [1] are those networks that are not physically connected by wires like Ethernet, giving the user a lot of mobility and convenience. Additionally, it saves money on the wires that would be needed to use a wired network and makes it simpler to move the base of the devices from one location to another by simply relocating the computer and wireless network card. A wired network facilitates point-to-point communication, or the transmission of data

between Access points (APs), which link to a public or private (enterprise) network, wireless clients used by users, and airwave data transmission are the four fundamental parts of wireless networks. Fig. 1 depicts the fundamental elements of wireless networking.

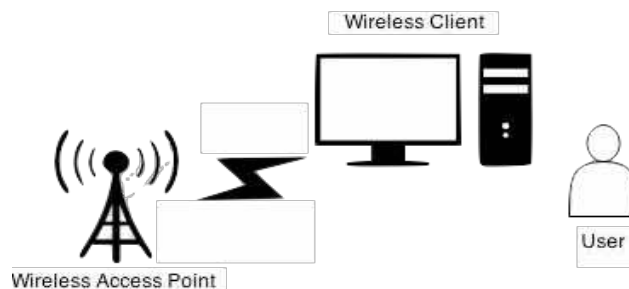


Fig. 1. Wireless Network Components [2]

Compared to wired networks, wireless networks produce a lot of sensitive data that is more susceptible to interception. Users of wireless networks may opt to use a variety of encryption techniques to reduce the danger because of this greatly increased risk. The secret to keeping data secure online in a Wi-Fi network is encryption. Information gets scrambled into a code when it is encrypted, making it unreadable to outsiders. Numerous encryption techniques have been developed because of the heightened likelihood of information compromise linked with Wi-Fi networks. However, widely used encryption techniques are known to contain flaws and are open to attack compromising confidentiality.[5]

Communication between nodes (users) and base station (BS) to other nodes should be managed carefully by means of an effective key management protocol to make secure communications throughout a wireless network. Using quantum cryptography, a new key distribution technique called quantum key distribution (QKD) enables the broadcast of a network key with complete confidentiality.[26] By utilizing the characteristics of quantum information, this technique resolves the issue of key distribution and offers a secure communication network between two users with unwavering security. The methodology for key distribution in wireless networks utilizing quantum cryptography and accompanying protocols is presented in this study. The study also explores additional strategies for maintaining the security and safety of wireless networks.[32]

2. RELATED WORKS

Any quantum key distribution (QKD) protocol aims to provide a shared secret key that can be used to communicate secretly between two remote parties across a common communication channel. The most important thing to remember in this situation is that the key generating protocol is demonstrably secure against any attack that an eavesdropper could launch.[24] The security of the protocol is guaranteed by the law of physics (or, more precisely, quantum mechanics), not just by the technological constraints that come with actual implementations. Therefore, one may be confident that the protocol will be secure for all time, not only until someone creates a ridiculously potent decryption device (or, to be precise, the protocol will be secure as long as quantum mechanics is not disproved).[9]

A quantum key distribution technique may generally be split into two parts: The quantum transmission phase, which takes up the first section, is where Alice and Bob send and/or measure quantum states. The second stage is the classical post-processing phase, where a pair of secure keys are created from the bit strings produced in the quantum phase.[28]

Some of the characteristics of 802.11 WLANs that give it the proper environment for incorporating QKD are as follows:

1. **Usage:** 802.11 WLANs are mainly used in office and campus infrastructure, which facilitates the deployment of QKD network with a high density of quantum apparatus if necessary.
2. **Capacity:** Terminals present in a 802.11 WLAN have more computational capacity and more energy as compared to those in cellular networks.
3. **Connection:** 802.11 WLANs are used to provide access to the Internet through an AP installed in an organization, which is beneficial for QKD integration.

To facilitate efficient authentication and management of keys between access point and client, along with user traffic control 802.11 networks employs Extensible Authentication Protocol (EAP) [11]. EAP provides an authentication framework, which will be used in the current work. The security of 802.11 WLANs is based on the WEP protocol, with enhanced security being provided by the MAC layer. As specified earlier WEP presents security problems in a wireless network. The MAC layer on the other hand defines two classes for security: Robust Security Network Association (RSNA) [10] and Transaction Security Network (TSN) [10].

Routing Protocol for Wireless Networks:

The directing conventions are utilized to track down the way among Source and Destination. There are various explanations behind planning and characterizing directing conventions for remote specially appointed networks. Directing conventions give essential job in the advanced correspondence networks.[26] Every single one of the steering conventions, has different design in contrast with others, thus every one of them, depending on boundaries which connected with the organization, exhibit fantastic execution. Chiefly directing conventions are separated into two classes: table-driven and on-request steering considering how to find the courses. In table driven directing conventions predictable what's more, forward-thinking steering data to all hubs is kept up with at every hub though in on-request directing the courses are made just when wanted by the source have. Moreover, these can likewise be comprehensively arranged into half and half and progressive. Every one of these four characterization are momentarily made sense of underneath alongside their models.[12]

A) Reactive Routing Protocol: The protocols find route on demand by flooding the network with Route Request (RREQ) packets to the nodes. In reactive routing protocols, the route is calculated only when a node needs to send data to an unknown destination. Thus, route discovery is initiated only when needed. Determine a route only when there is data to transfer. These protocols have longer delay and low routing overhead. e.g., AODV, AOMDV.[15]

The routing protocol known as AOMDV: Ad hoc On-Demand Multipath Distance Vector (AMODV) Routing [11] is used in MANETs and other wireless ad hoc networks. In comparison to DSDV [12], which reduces the number of necessary broadcasts by constructing routes on demand, it is an improvement. Nodes that do not follow the chosen path cannot share routing tables or contribute to the upkeep of routing information. By broadcasting an RREQ packet to its neighbors, a source node starts a path discovery process to find the other intermediate nodes (and the destination).[11]

AODV: Ad hoc On-Demand Distance Vector (AODV) Routing [11] is used in MANETs and other wireless ad hoc networks. In comparison to DSDV [12], which reduces the number of necessary broadcasts by constructing routes on demand, it is an improvement. Nodes that do not follow the chosen path cannot share routing tables or contribute to the upkeep of routing information. By broadcasting an RREQ packet to its neighbors, a source node starts a path discovery process to find the other intermediate nodes (and the destination).

DSR: Ad-Hoc Wireless Network's Dynamic Source Routing(DSR) [13]. This style of source routing is on demand. Every host keeps a route cache with all the routes it has discovered. The Route Discovery and Route Maintenance stages make up the

protocol's main sections. A mobile node first checks its route cache to see if it has a route to a certain destination when it has a packet to send to that location. It will take this route if the route is still valid. In the absence of a route, the node starts the process of route discovery by broadcasting a Route Request packet.

B) The pro-active routing protocol: It discovers the way in advance, and each node tries to maintain an up-to-date topological map of the complete network and continuously assess the routes. There is no delay in route discovery because the route is already open when a node wants to forward a packet. It is an effort to preserve accurate, consistent routing data. In order to retain a consistent

perspective when the network topology changes, the protocol reacts by propagating updates throughout the network. The routing overhead with this kind of table-driven method is substantial. e.g. – DSDV [16]

DSDV: Bellman-Ford-based Highly Dynamic Destination-Sequence Distance-Vector Routing (DSDV) for Mobile Computers. According to the number of hops needed to reach each destination, each mobile node keeps a routing table. Updates to the routing table are sent out periodically. A sequence number generated by the destination node is attached to each entry. Every node broadcasts its distance vector on a regular basis. One hop is the maximum for broadcast.

C) Hybrid Protocols: A third category of routing algorithms is called hybrid routing [10]. It combines being proactive and being reactive. It benefits from both. To address their drawbacks, it combines the advantages of proactive and reactive routing systems. Hybrid MANET routing protocols typically make use of hierarchical network architectures. All routers will first set up specific proactive routes and begin processing data. After that, a flooding of several RR packets will enable an on-demand scenario to function. One of the hybrid forms of routing protocols is the Zone Routing Protocol (ZRP) [17].

D) Hybrid Protocols: The hierarchical level at which a node is located determines whether proactive or reactive routing should be used. The foundation of hierarchical routing is the division of nodes into groups and the assignment of distinct functionality to nodes both inside and outside of a group. Using CBRP [18] and FSR [19] as examples.

Security Issues in Wireless Networks:

Since the transmission in this case uses the medium of air rather than cables, wireless networks do not guarantee quality of service during transmission, and the likelihood of intrusion into such networks is very high. Therefore, it not only needs to prevent unauthorized users from entering the network, but it also needs to secure the transmission of users' confidential data. The following [2] are the general security

concerns for wireless networks:

- 1) **Confidentiality:** Since the information being transmitted across the network is encrypted in transit to ensure that only the intended recipient can read it, authentication of the receiver is also necessary. The receiver will be given the key to decrypt the information they receive.
- 2) **Integrity:** Wireless networks are vulnerable to assaults that could compromise the integrity of the data. The integrity prevention techniques utilized are comparable to those in wired networks.
- 3) **Availability:** Denial of service attacks are possible on wireless networks. Network accessibility can be restricted using radio jamming. Another attack called, battery exhaustion attack is also arisen where unauthorized users repeatedly send messages or data to connect devices and hence runs down the device's battery.
- 4) **Eavesdropping and Authentication:** Since wireless networks are broadcast, as was already noted, there are a greater number of access points that can be utilized to connect to the network. This eavesdropping must be stopped.
- 5) **Blue Snaring or Blue jacking:** These are the Bluetooth attacks carried out to alter or steal data.

Classical Cryptography Techniques

1. Symmetric Key Cryptography:

Symmetric key distribution-based cryptosystems employ the same key for both encryption and decryption. It is also referred to as secret key cryptography. Only if the symmetric keys are distributed beforehand to each pair of interacting systems can a secure communication channel in key management be accomplished. Fig 2. depicts the symmetric key cryptography procedure.

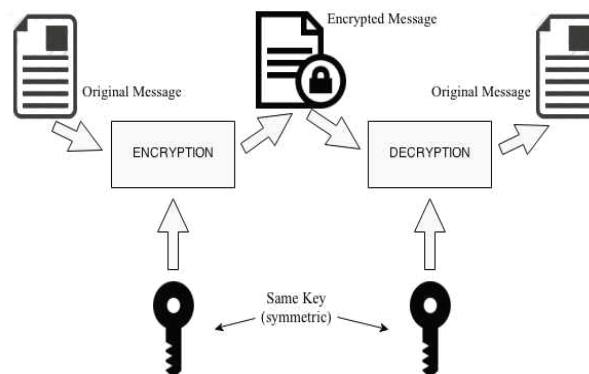


Fig. 2. Symmetric Key Cryptography [4]

2. Asymmetric Key Cryptography

A public key system, which consists of two components—a private key that is kept secret and a public key that is disseminated over the network—is used by cryptosystems that employ asymmetric key distribution.

Using the recipient's public key, the sender encrypts the message. To decode the communication, the receiver uses its private key. The private key is never in transit in this distribution, making it less susceptible to security problems. Fig. 3. depicts the asymmetric key cryptography process in which a device, such as a laptop, seeks to connect to an unsecured network and might capture the personal information of a user using that network.

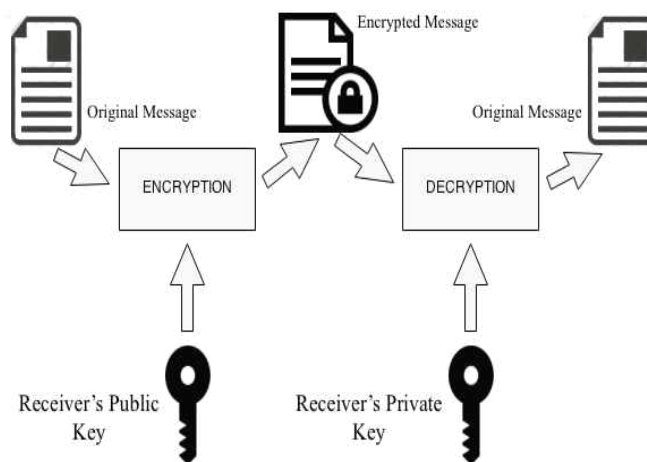


Fig. 3. Asymmetric Key Cryptography [2]

3. QUANTUM KEY DISTRIBUTION PROTOCOLS

The goal of any quantum key distribution (QKD) protocol is to generate a shared secret key between two distant parties over a public communication channel. The crucial point here is that the key generating protocol is provably secure against any possible attack that an eavesdropper can perform. It is the law of physics (or, in fact, quantum mechanics) that guarantees the security of the protocol, not only the technical limitations that exist in practical implementations.

Six-State Protocol (SSP)

The SSP protocol follows the identical procedures as the BB84 protocol, with the exception that it permits the encoding of the bits needed for inter-entity communication using three orthogonal bases. As a result, the bits are represented by six states [7].

Quantum Node Networks

A quantum node can be used to combat the quantum decoherence of the signal along

the quantum channel by actively performing quantum operations on the travelling photons. One way to do this suggested by Cirac et al. [JC98] necessitates the use of quantum entanglement sources, quantum memories and entanglement purifications techniques which result in perfect entangled states subsequently stored in a portion of the quantum channel.[6] These nodes are Quantum Repeaters: they chain the stored states together and thus obtain perfect end-to-end entanglement useable over arbitrarily long distances. However, quantum repeaters exist only theoretically: they await the development of fully fledged quantum computing capabilities to become practical. A less elaborate quantum node, the Quantum Relay proposed by Collins et al. [DC03+ doesn't require a quantum memory and is hence feasible but technologically difficult. However, quantum relays do not extend the distance over which a quantum signal can be sent so will not be suitable for a practical QKD network.[21]

Trusted Relay Networks

A relay node is trusted implicitly to forward on the quantum signal without eavesdropping or tampering with it. To do this in a QKD network, local keys are generated over QKD links and stored securely in trusted nodes at each end of the links. (The nodes are effectively a mini-Alice and mini-Bob performing their own little QKD protocol run, independently of any other messages being passed through the network.) When a real Alice and real Bob want to do a QKD protocol run, a chain of trusted relays and their intermediate quantum links is created to connect them together: this is a QKD path. Alice and Bob's quantum key is treated as a message and encrypted via a onetime pad using a local key stored at a trusted node: it travels "hop by hop" between each node on the QKD path and is decrypted and re-encrypted at each node using a new key from the node key store.

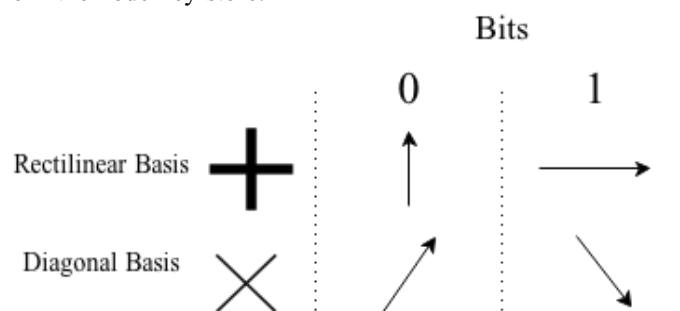


Fig. 4. Photon Polarization using Bases [9]

After receiving all the photons, the receiver now communicates with the sender using the public channel for key sifting.

Using Classical channel (Key Sifting):

Receiver informs the Sender what bases he used to measure the photons and Sender responds by saying if it matched the bases used.

4. PROPOSED PROTOCOL

In order to overcome the security issues of key distribution this paper employs QKD protocol, which is an improved version of BB84 and Fig. 6 shows the QKD protocol being implemented in the proposed 4-way handshake protocol.

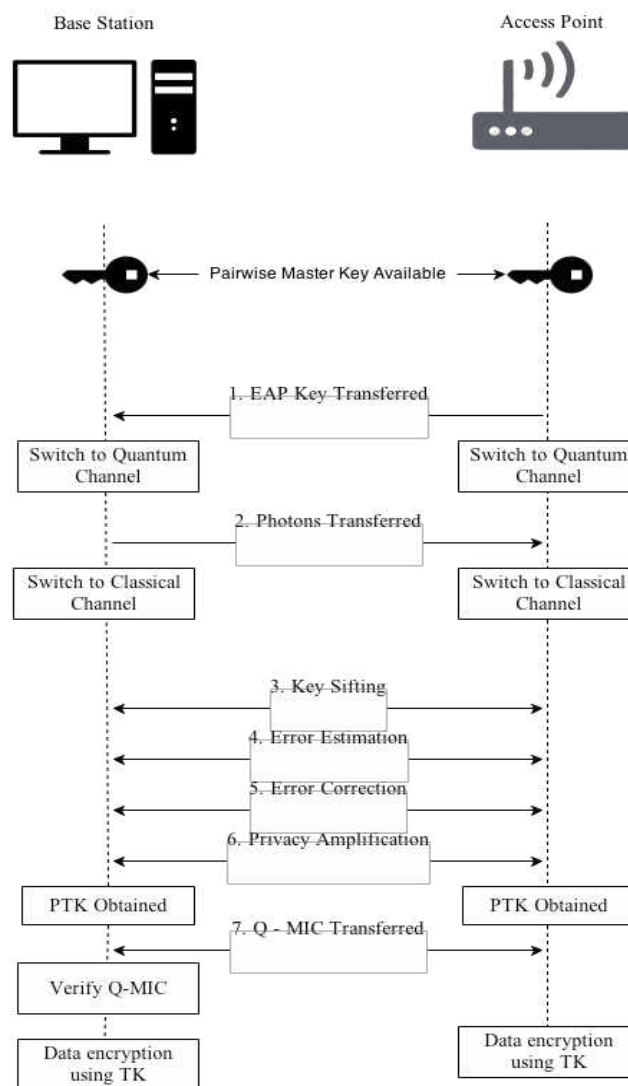


Fig. 5. Proposed Handshaking Protocol [6]

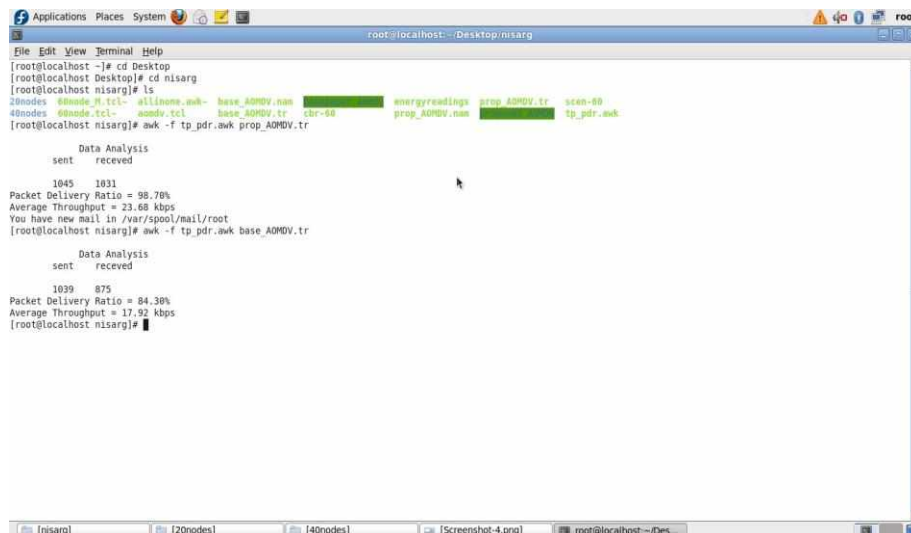
The steps involved in the proposed protocol are as follows:

1. Initially the PMK is shared between the BS and the AP. Then the transmission process is switched to the quantum channel.
2. BS sends all the polarized photon to the AP using bases at random. As soon as the transmission of photons is finished, the channel is switched to classical.
3. The next 3 stages of QKD are applied to remove all the error and obtain the final encryption key.

Quantum transmission must convey enough photons in order to strengthen the quantum key and provide safe communication. The quantum key (384 bits) is reduced to the required number of bits to create the PTK key(256 bits). The standard 4-way handshake procedure is then performed after obtaining the PTK key, during which the MIC is derived for mutual authentication. The initial half of equal-length bits in PMK and an XOR operation are used to operate the MIC. Quantum MIC is the name of this MIC (Q-MIC) between the two parties, Q-MIC is transmitted and confirmed. After that, communication and encryption are started using the temporal key. The Nounce values used in the initial handshake are not necessary because the QKD protocol is being used. The next section lists additional benefits of using this procedure.[14]

Implementation Result:

I have implemented the proposed Quantum cryptography with AOMDV network routing protocol and compared its result with the AOMDV without using Quantum cryptography.[18]



PDR delivery ratio: It is the ratio of number of packets received at the destination to the number of packets sent from the source.

Graph 1. Packet Delivery Ratio for different number of nodes

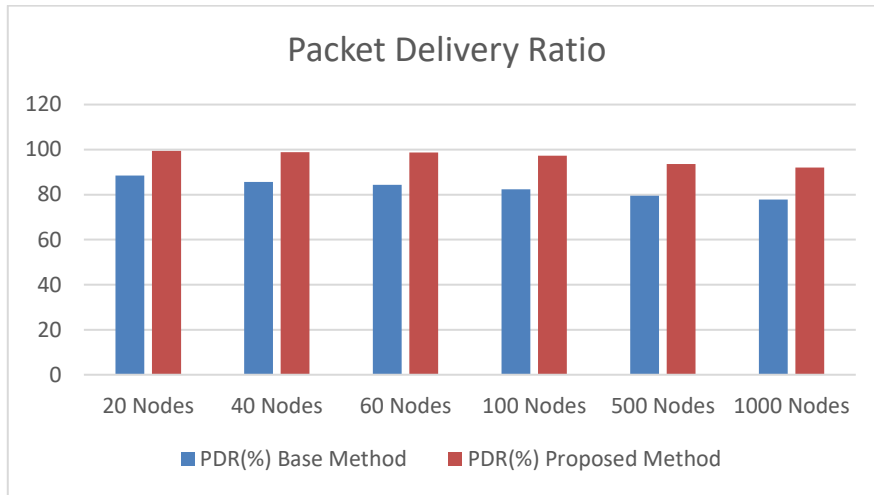


Table 1. Packet Delivery Ratio for different number of nodes

Nodes	PDR(%) – Base	PDR(%) – Proposed
20	88.54	99.42
40	85.63	98.89
60	84.30	98.70
100	82.32	97.27
500	79.58	93.62
1000	77.83	92.10

Throughput: It is the ratio of the total data reaches a receiver from the sender.

Graph 2. Packet Delivery Ratio for different number of nodes

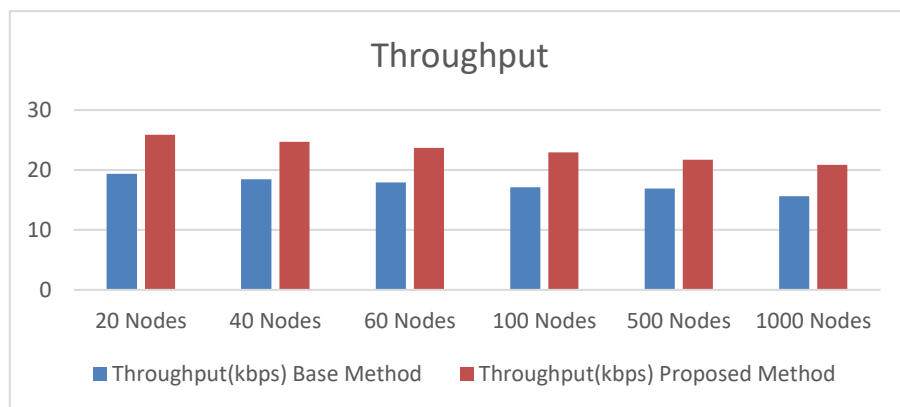


Table 2. Throughput for different number of nodes

Nodes	Throughput (kbps) – Base	Throughput (kbps) – Proposed
20	19.37	25.87
40	18.48	24.73
60	17.92	23.68
100	17.10	22.93
500	16.89	21.74
1000	15.61	20.86

Advantages of quantum cryptography

Quantum cryptography is still a long way from being implementation is far from the theoretical technique. But still there is no debate that quantum cryptography is a true breakthrough in network security. Some of the advantages of applying quantum cryptography in wireless networks are as follow [18]:

- Quantum Cryptography approach doesn't depend on mathematics models but instead is based on physics principles for decoding an encoded data making it virtually non-hack able and requires few resources for its maintenance.[8]
- QKD is found useful for different categories of wireless networks as a means for connecting devices to different access points in proximity. Hence making it more

efficient for communication with fewer cables.

5. CONCLUSION AND FUTURE WORK

The main goal of this research work is to show a method to improve the security aspect of WLANs. It has been shown that the integration of Quantum Cryptography in Wireless Networks has great prospective in terms of better network security and performance.

In Future, we can use quantum cryptography with other routing protocols as well.

6. REFERENCES

1. Matthias J Kannwischer, Aymeric Genet, Denis Butin, Juliane Kr ^ amer, and Johannes Buchmann. Differential power analysis of xmss and sphincs. In International Workshop on Constructive Side-Channel Analysis and Secure Design, pages 168–188. Springer, 2018.
2. Laurent Castelnovi, Ange Martinelli, and Thomas Prest. Grafting trees: a fault attack against the sphincs framework. In International Conference on Post-Quantum Cryptography, pages 165–184. Springer, 2018.
3. Dorian Amiet, Lukas Leuenberger, Andreas Curiger, and Paul Zbinden. Fpga-based sphincs+ implementations: Mind the glitch. In 2020 23rd Euromicro Conference on Digital System Design (DSD), pages 229–237. IEEE, 2020.
4. Brice Colombier, Vlad-Florin Dragoi, Pierre-Louis Cayrel, and Vincent Grosso. Message-recovery profiled side-channel attack on the classic mceliece cryptosystem. Cryptology ePrint Archive, Paper 2022/125,2022. <https://eprint.iacr.org/2022/125>.
5. Joppe W Bos, Marc Gourjon, Joost Renes, Tobias Schneider, and Christine van Vredendaal. Masking kyber: First-and higher-order implementations. IACR Transactions on Cryptographic Hardware and Embedded Systems, pages 173–214, 2021.
6. Prasanna Ravi, Romain Poussier, Shivam Bhasin, and Anupam Chattopadhyay. On configurable sca countermeasures against single trace attacks for the ntt. In International Conference on Security, Privacy, and Applied Cryptography Engineering, pages 123–146. Springer, 2020.
7. Prasanna Ravi, Sujoy Sinha Roy, Anupam Chattopadhyay, and Shivam Bhasin. Generic side-channel attacks on cca-secure lattice-based pke and kems. IACR Transactions on Cryptographic Hardware and Embedded Systems, pages 307–335, 2020.
8. Prasanna Ravi, Anupam Chattopadhyay, and Anubhab Baksi. Sidechannel and fault-injection attacks over lattice-based post-quantum schemes (kyber, dilithium): Survey and new results. Cryptology ePrint Archive, 2022.

9. NIST. Post-Quantum Crypto Project - Round 3 Submissions. <https://csrc.nist.gov/Projects/post-quantum-cryptography/round-3-submissions/>, 2021.
10. Dustin Moody, Gorjan Alagic, Daniel Apon, David Cooper, Quynh Dang, John Kelsey, Yi-Kai Liu, Carl Miller, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, and Jacob Alperin-Sheriff. Status report on the second round of the nist post-quantum cryptography standardization process, 2020-07-22 2020.
11. Kyungbae Jang, Anubhab Baksi, Jakub Breier, Hwajeong Seo, and Anupam Chattopadhyay. Quantum implementation and analysis of default. Cryptology ePrint Archive, Paper 2022/647, 2022. <https://eprint.iacr.org/2022/647>.
12. Zhenyu Huang and Siwei Sun. Synthesizing quantum circuits of AES with lower t-depth and less qubits. IACR Cryptol. ePrint Arch., page 620, 2022.
13. Kyungbae Jang, Anubhab Baksi, Gyeongju Song, Hyunji Kim, Hwajeong Seo, and Anupam Chattopadhyay. Quantum analysis of AES. IACR Cryptol. ePrint Arch., page 683, 2022.
14. Thomas Haner, Samuel Jaques, Michael Naehrig, Martin Roetteler, and Mathias Soeken. Improved quantum circuits for elliptic curve discrete logarithms. In Jintai Ding and Jean-Pierre Tillich, editors, Post-Quantum Cryptography, pages 425–444, Cham, 2020. Springer International Publishing.
15. Craig Gidney and Martin Eker. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. Quantum, 5:433, April 2021.
16. Andre Chailloux, Maria Naya-Plasencia, and Andre Schrottenloher. An efficient quantum collision search algorithm and implications on symmetric cryptography. Cryptology ePrint Archive, Paper 2017/847, 2017. <https://eprint.iacr.org/2017/847>.
17. Frank et al Arute. Quantum supremacy using a programmable superconducting processor. Nature, 574(7779):505–510, 2019.
18. Wallden, P. and E. Kashefi (2019). "Cyber security in the quantum era." Communications of the ACM 62(4): 120-120.
19. Patrick Grofig, Martin Härterich, Isabelle Hang, Florian Kerschbaum, Mathias Kohler, Andreas Schaad, Axel Schröpfer, and Walter Tighzert. Experiences and observations on the industrial implementation of a system to search over outsourced encrypted data. In Stefan Katzenbeisser, Volkmar Lotz, and Edgar R. Weippl, editors, Sicherheit, volume 228 of LNI, pages 115–125. GI, 2014.
20. Shweta Singh, Anil Kr. Ahlawat, and Arun Kr. Tripathi: "Mechanizing Wireless LAN (WLAN) using Compression Techniques", in Free Journal International Journal of Engineering, Applied Science and Technology (IJEAST), ISSN No: 2455-2143, March-April 2016.
21. Zhou, T., et al. (2018). "Quantum cryptography for the future internet and the security analysis." Security and Communication Networks 2018.
22. J. Shen, T. Zhou, X. Chen, J. Li, and W. Susilo, "Anonymous and Traceable Group Data Sharing in Cloud Computing," IEEE Transactions on Information Forensics and Security, vol. 13, no. 4, pp. 912–925, 2018.
23. J. Li, Y. Zhang, X. Chen, and Y. Xiang, "Secure attribute-based data sharing for resource-limited users in cloud computing," Computers & Security, 2017.

24. J. Shen, J. Shen, X. Chen, X. Huang, and W. Susilo, "An efficient public auditing protocol with novel dynamic structure for cloud data," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 2402–2415, 2017.
25. J. Shen, T. Miao, Q. Liu, S. Ji, C. Wang, and D. Liu, "S-SurF: An Enhanced Secure Bulk Data Dissemination in Wireless Sensor Networks," in *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, vol. 10656 of *Lecture Notes in Computer Science*, pp. 395–408, Springer International Publishing, Cham, 2017.
26. P. Li, J. Li, Z. Huang, C.-Z. Gao, W.-B. Chen, and K. Chen, "Privacy-preserving outsourced classification in cloud computing," *Cluster Computing*, pp. 1–10, 2017.
27. C. H. Bennett, G. Brassard, C. Crépeau, and M.-H. Skubiszewska, "Practical quantum oblivious transfer," in *Annual International Cryptology Conference*, pp. 351–366, Springer.
28. S. Winkler and J. Wullschleger, "On the efficiency of classical and quantum oblivious transfer reductions," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: Preface, vol. 6223, pp. 707–723, 2010.
29. A. Chailloux, I. Kerenidis, and J. Sikora, "Lower bounds for quantum oblivious transfer," *Quantum Information & Computation*, vol. 13, no. 1-2, pp. 0158–0177, 2013.
30. J. Shen, T. Zhou, D. He, Y. Zhang, X. Sun, and Y. Xiang, "Block design-based key agreement for group data sharing in cloud computing," *IEEE Transactions on Dependable and Secure Computing*, vol. PP, no. 99, 2017.
31. J. Shen, D. Liu, J. Shen, Q. Liu, and X. Sun, "A secure cloud-assisted urban data sharing framework for ubiquitous-cities," *Pervasive and Mobile Computing*, vol. 41, pp. 219–230, 2017.
32. B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, "Hiding bits in bell states," *Physical Review Letters*, vol. 86, no. 25, pp. 5807–5810, 2001.