

jauti2025.web.ua.pt

PROCEEDINGS

JAUTI'25

14ª Jornadas de Aplicaciones
y Usabilidad de los Medios
Interactivos

14th Iberoamerican Conference
on Applications and Usability
for Interactive Media

28
NOVEMBER

UNIVERSIDADE DE AVEIRO
Portugal



universidade de aveiro
theoria poiesis praxis



DigiMedia
digital media and interaction
research centre



RedAUTI

TITLE

Proceedings of the 14th Iberoamerican Conference on Applications and Usability for Interactive Media - jAUTI'2025

EDITORS

Pedro Almeida - Digimedia – University of Aveiro, Portugal
Ana Velhinho - Digimedia – University of Aveiro, Portugal
Telmo Silva - Digimedia – University of Aveiro, Portugal
Jorge Ferraz de Abreu - Digimedia – University of Aveiro, Portugal
Maria José Abasolo Guerrero - University of La Plata, Argentina

AUTHORS

Abílio Almeida, Adriana Silvia Fachal, Alex Chiguano, Ana Camacho, Ana I. Lillebø, Ana Isabel Veloso, Ana Martins, Ana Patricia Oliveira, Ana Teixeira, Ana Teresa Bessa, Ana Velhinho, Ana Vitória Maciel de Lima Pessoa, Anabela G. Silva, Anabela Silva, Andrea Miranda, Anna Broberg, Arthur Henrique da Silva, Asta Hiippala, Bárbara Silva, Bryan Carrillo, Carlos Queirós, Hélder Caixinha, Carminda Lomba, Catarina Sousa, Ceci Diehl, Cecilia Verónica Sanz, Clarissa Evelin Pereira, Cláudia Pedro Ortet, Daniel Alves, Daniel Crespo, David Badajoz, Dionísia Laranjeiro, Enrickson Varsori, Fabio Henrique Ferreira de Albuquerque, Fabrício Marques de Sousa, Filomena Sobral, Francisco José Perales, Gabriel Aguiar, Galo Cornejo, Giulia Pintan de Souza Sant'Ana, Gonçalo Santinha, Gonzalo Olmedo, Graciela Santos, Gustavo Javier Astudillo, Helena Gomes, Isabel Costa, Joan Josep Ordoñez, Joao Marcelo Alves Macedo, Jónatas Roque, Jorge Abreu, Jose Maria Buades, José Ramón Beltrán, Liliana Vale Costa, Luís Pedro María José Abásolo Guerrero, Mariana Pinho, Maristela de Freitas Riquelme, Martín Fernando Puricell, Mayra Comina, Mirian Solis, Mónica Aresta, Nancy Paredes, Nelson Benavides, Nelson Rocha, Nídia Salomé Morais, Óscar Ribeiro, Pedro Almeida, Pedro Beça, Pedro Rito, Raquel Magalhães, Rita Oliveira, Roberto Ariel Guerrero, Rodrigo Graça, Rui Raposo, Sandra Baldassarri, Sofia Chicaiza, Sofia Ribeiro, Sónia Ferreira, Sónia Samagaio, Sudarshan Kadam, Telmo Silva, Teresa Gouveia, Teresa Marques, Valdecir Becker, Víctor Zurita, Wesley Alves

EDITORIAL DESIGN

Ana Velhinho

PUBLISHER

Universidade de Aveiro
1st Edition – April 2026
ISBN 978-989-9253-61-2
DOI <https://doi.org/10.48528/pazy-q864>

The sole responsibility for the content of this publication lies with the authors. © Authors.
This work is licensed under a Creative Commons Attribution 4.0 International License.

jauti2025.web.ua.pt

PROCEEDINGS

JAUTI'25

14ª Jornadas de Aplicaciones
y Usabilidad de los Medios
Interactivos

14th Iberoamerican Conference
on Applications and Usability
for Interactive Media

28
NOVEMBER

UNIVERSIDADE DE AVEIRO
Portugal



COMMITTEE

PROGRAM CHAIRS | ORGANIZING COMMITTEE

Pedro Almeida

Digimedia – University of Aveiro, Portugal

Ana Velhinho

Digimedia – University of Aveiro, Portugal

Telmo Silva

Digimedia – University of Aveiro, Portugal

Jorge Ferraz de Abreu

Digimedia – University of Aveiro, Portugal

Maria Jose Abasolo Guerrero

University of La Plata, Argentina

SCIENTIFIC COMMITTEE

Alan Guedes - University of Reading, UK

Alcina Prata - Polytechnic Institute of Setúbal, Portugal

Ana Velhinho - Digimedia – University of Aveiro, Portugal

Anelise Jantsch - Federal University of Rio Grande do Sul, Brazil

Angel Garcia Crespo - Carlos III University, Spain.

Antoni Bibiloni LTIM - University of Balearic Islands, Spain

Antoni Oliver LTIM - University of Balearic Islands, Spain

Armando De Giusti - III-LIDI, National University of La Plata, Argentina

Beatriz Sainz de Abajo - University of Valladolid, Spain

Carlos de Castro Lozano - University of Córdoba, Spain

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

Daniel Gambaro - University of São Paulo, Brazil

Douglas Paredes Marquina - University of Los Andes, Venezuela

Fernanda Chocron Miranda - 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 Montero Simarro - University of Castilla-La Mancha, Spain

Francisco Perales López - UGIV-IA University of Balearic Islands, Spain

Gabriel Fernandez - Ramon Llull University, Barcelona, Spain

Gonzalo Fernando Olmedo Cifuentes - Polytechnic School of the Army ESPE, Ecuador

Guido Lemos - LAVID – Federal University of Paraíba, Brasil

Gustavo Rossi - National University of La Plata, Argentina

Israel González Carrasco - University Carlos III, Spain

Ivan Bernal - National Polytechnic School, Ecuador

Joaquín Danilo PINA AMARGÓS - CUJAE, Cuba

Jordi Belda - Polytechnic University of Valencia, Spain

Jorge Abreu - Digimedia – University of Aveiro, Portugal

Jorge Eduardo Guaman Jaramillo - Universidad Técnica Particular de Loja, Ecuador

José Luis Arciniegas Herrera - University of Cauca, Colombia

Josemar Rodrigues de Souza - University of Bahia State, Brazil

Jose Maria Buades Rubio - UGIV-IA University of Balearic Islands, Spain

Juan Carlos Torres - Universidad Técnica Particular de Loja, Ecuador

Larissa Coto Valdeperas - Universidad de Costa Rica

Luis Enrique Martínez Martínez - University of Alicante, Spain

Luís Pedro - University of Aveiro, Portugal

Manuel Gonzalez Hidalgo - UGIV-IA University of Balearic Islands, Spain

Marcelo Fernandes de Sousa - Higher Education Institute of Paraiba, Brasil

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

Pablo Rodríguez Bocca - University of the Republic, Uruguay

Patrícia Oliveira - Digimedia – University of Aveiro, Portugal
Pedro Almeida - Digimedia – University of Aveiro, Portugal
Raisa Socorro Llanes - CUJAE, Cuba
Raoni Kulesza - Federal University of Paraíba, Brazil
Raphael Irerê - Catholic University of Brasília, 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, Brasil
Rui Manuel Raposo - University of Aveiro, Portugal
Sandra Baldassarri - University of Zaragoza, Spain
Sandra Casas - National University of Southern Patagonia, Argentina
Tatiana Tavares - Federal University of Pelotas, Brasil
Telmo Silva - Digimedia – University of Aveiro, Portugal
Tiago Maritan - Federal University of Paraíba, Brasil
Vagner Beserra - Universidad of Tarapacá, Chile
Valdecir Becker - Federal University of Paraíba, Brasil

FOREWARD

We welcome you to the proceedings book of the 14th Iberoamerican Conference on Applications and Usability for Interactive Media - jAUTI'2025. This conference is a joint organization of the Social Interactive Viewing Experiences (iVX) group of the DigiMedia Research Unit of the University of Aveiro and RedAUTI (thematic network on Applications and Usability of Interactive Digital Television).

As new challenges arise for RedAUTI and its scientific lens and scope, the conference chairs recognize the importance of holding this event to reinforce a unique Ibero-American network of researchers in the field of Digital and Interactive TV and update the conference name to the upcoming opportunities of the ecosystem, from the focus on Interactive TV Applications to richer mediated experiences within the Interactive Media field.

This conference edition (held on November 28th, 2025) took place in a hybrid format at the University of Aveiro, Portugal. Nevertheless, it brought together researchers from several universities, including Europe and the Americas, as well as industry professionals, to share their research and innovative products. This book collects a set of 26 papers organized in 8 thematic sessions, two of them integrated in a Workshop from the R&D project Polariscope on “Digital Storytelling and AI in Co-Creation processes”: Generative AI and Video Experiences, Advanced Services and TV 3.0, Usability and UX Evaluations, Digital Accessibility, EEG and Affective Computing, Interactive Video Experiences, Digital Storytelling for Student Engagement, and Applications for Digital Storytelling.

The scientific paper presentations were followed by engaging discussions, moderated by the chairs of each session, and a hands-on activity within the scope of the Polariscope Workshop, focused on collaborative storytelling through media content.

We wish to thank the dedication of all the members of the Scientific Committee, as well as those who helped us organise the conference.

Our warm regards,
Pedro Almeida, Ana Velhinho, Telmo Silva,
Jorge Ferraz de Abreu, Maria José Abasolo Guerrero

TABLE OF CONTENTS

Generative AI and Video Experiences

- 014 | ***Navigating the Dual Challenge: A Strategic Framework for Ethical and Pedagogical Integration of Generative AI in Higher Education***
Nídia Salomé Morais, Teresa Gouveia, Rui Raposo, Pedro Rito, Catarina Sousa, Sónia Ferreira, Filomena Sobral, Helena Gomes
- 028 | ***Aspetos audiovisuais na construção de vídeos de exercício físico para pessoas mais velhas com dor crónica***
Ana Teresa Bessa, Telmo Silva, Anabela Silva
- 037 | ***Documenting the Production Procedures of an In-Motion 360-Degree Video for Senior Virtual Cycling Tourism Experiences***
Fabrício Marques, Cláudia Pedro Ortet, Liliana Vale Costa, Ana Isabel Veloso

Advanced Services and TV 3.0

- 046 | ***Rede Pública: Portal de Integração para Conteúdos Audiovisuais e Serviços Digitais na TV 3.0 Brasileira***
Valdecir Becker, Wesley Alves da Silva, Maristela Riquelme, Clarissa Pereira, Bárbara Silva, Gabriel Aguiar
- 054 | ***TV Senado na Era da TV 3.0: Uma Proposta de Experiência do Usuário para o Acesso à Informação Legislativa***
Maristela de Freitas Riquelme, Clarissa Evelin Pereira, Wesley Alves da Silva, Valdecir Becker
- 061 | ***Automatic System for Reception, Processing, and Subtitling of ISDB-Tb One-Seg Signals for Hearing Accessibility***
Gonzalo Olmedo, Victor Zurita, Bryan Castillo, Nelson Benavides

Usability and UX Evaluations

- 070 | ***Development of a Controlled Robotic Prototype for Fatigue Recognition Through Microexpression***
Nancy Paredes, Alex Chiguano, Sofia Chicaiza, Mayra Comina
- 077 | ***Estimulación Cognitiva mediante Juegos Serios: Aceptación de la Televisión Digital y Análisis Comparativo de la Respuesta Emocional Multidispositivo en Adultos Mayores***
Ana Camacho , María José Abasolo, Rita Oliveira , Mirian Solis, Galo Cornejo
- 088 | ***Framework liviano y abierto para el desarrollo de Presentaciones interactivas de Realidad Extendida***
Martín Fernando Puricell, María José Abásolo Guerrero, Roberto Ariel Guerrero
- 095 | ***Evaluating the Usability and Effectiveness of an Interactive Platform for Stakeholder Engagement and Empowerment***
Enrickson Varsori, Mariana Pinho, Dionísia Laranjeiro, Daniel Crespo, Asta Hiippala, Anna Broberg, Ana I. Lillebø
- 103 | ***Enhancing Accessibility and Usability in DanceMove: Expert Evaluation of a Foot-Interaction Game for Older Adults***
Ceci Diehl, Telmo Silva, Jorge Ferraz, Gonçalo Santinha, Óscar Ribeiro, Ana Martins, Nelson Rocha, Anabela G. Silva

Digital Accessibility

- 112 | ***Revisión Sistemática sobre Entornos Virtuales 3D en la Educación de Personas con Discapacidad Auditiva***
Adriana Adriana Fachal, Maria José Abásolo, Cecilia Verónica Sanz
- 121 | ***ISDB-T One-Seg Receiver with Automatic Integration of Ecuadorian Sign Language (LSEC) Interpretation Using Neural Networks***
Gonzalo Olmedo, Bryan Carrillo, Víctor Zurita , Nancy Paredes
- 130 | ***MEO4ALL: Enhancing Interactive Television Accessibility through Voice Interaction and Inclusive Interface Redesign***
Rita Oliveira, Ana Patrícia Oliveira, Daniel Alves
- 138 | ***Open Government, Accessibility, and Limitations to Citizen Participation: The Brazilian Case***
João Marcelo Alves Macedo, Fábio Henrique Ferreira de Albuquerque, Valdecir Becker

EEG and Affective Computing

- 150 | ***Dataset characterization for user emotional analysis***
Joan Josep Ordoñez, Jose Maria Buades , Francisco José Perales, Sandra Baldassarri, José Ramón Beltrán, Gustavo Javier Astudillo
- 159 | ***EEG e UX: Como a Análise das Ondas Cerebrais Pode Contribuir para a Área de Experiência do Usuário***
Wesley Alves da Silva, Maristela de Freitas Riquelme, Giulia Pintan Souza de Sant'Ana, Arthur Henrique da Silva, Ana Vitória Maciel de Lima Pessoa, Valdecir Becker

Interactive Video Experiences

- 168 | *Interactive Shoppable Video Strategies for Specific Markets: A Focus on Portugal and Spain*

Sudarshan Kadam, David Badajoz, Telmo Silva, Jorge Abreu

- 178 | *Pausa para pensar: Análisis de la Demanda Cognitiva y la Intencionalidad Didáctica en Videos Interactivos.*

Andrea Miranda , Graciela Santos , María José Abásolo Guerrero

POLARISCOPE WORKSHOP: "Digital Storytelling and AI in Co-Creation processes"

Digital Storytelling for Student Engagement

- 193 | *Para Além dos Pontos e Distintivos: Desenhar a Motivação Intrínseca nas Ferramentas Digitais de SEL*

Raquel Magalhães

- 201 | *Game Engagement and Environmental Literacy: Development and Validation of a Motivational Model in Digital Game Contexts*

Mariana Pinho, Enrickson Varsori, Abílio Almeida

210 | ***Digital to Physical Storytelling: From Digital Teacher Narratives to Tangible Student Engagement***

Sofia Ribeiro, Pedro Beça, Mónica Aresta

220 | ***O Storytelling Digital como Estratégia de Cocriação e Envolvimento da Comunidade Escolar***

Ana Velhinho, Carminda Lomba, Sónia Samagaio, Teresa Marques, Pedro Almeida, Luís Pedro

Applications for Digital Storytelling

232 | ***A Framework for AI-Assisted Storytelling from Collective Multimedia Records***

Rodrigo Graça, Pedro Almeida, Telmo Silva, Rui Raposo, Ana Velhinho, Ana Teixeira

245 | ***Do Facto ao Acontecimento: proposta de arquitetura narrativa interativa para o jornalismo digital***

Carlos Queirós, Hélder Caixinha

255 | ***Personalized Multimedia Story Viewing Experience on the Polariscope Platform***

Jónatas Roque, Pedro Almeida, Telmo Silva, Ana Velhinho

Generative AI and Video Experiences





Navigating the Dual Challenge: A Strategic Framework for Ethical and Pedagogical Integration of Generative AI in Higher Education

Nídia Salomé Morais¹[0000-0001-5680-4778], Teresa Gouveia¹[0000-0002-2519-6535],
Rui Raposo²[0000-0002-5275-6200], Pedro Rito¹[0000-0002-4507-0323], Catarina Sousa¹[0000-0001-7706-6225],
Sónia Ferreira¹[0000-0003-0664-1039], Filomena Sobral³[0000-0003-3679-7894], and
Helena Gomes⁴[0000-0001-7517-891X]

¹ Viseu School of Education, CI&DEI - Centre for Studies in Education and Innovation, Polytechnic Institute of Viseu, Viseu 3504-510, Portugal

² Department of Communication and Art, DigiMedia. University of Aveiro, 3810-193 Aveiro, Portugal

³ Viseu School of Education, Centre for Communication and Culture (CECC) & CI&DEI - Centre for Studies in Education and Innovation, Portugal

⁴ Viseu School of Education, Center for Research and Development in Mathematics and Applications (CIDMA), Portugal

salome@esev.ipv.pt, tgouveia@esev.ipv.pt, raposo@ua.pt,
rito@esev.ipv.pt, csousa@esev.ipv.pt,
sonia.ferreira@esev.ipv.pt, filomena@esev.ipv.pt,
hgomes@esev.ipv.pt

Abstract. The exponentially rapid ascent of Generative Artificial Intelligence (GenAI) and merging into multiple levels and niches of the higher education ecosystem has created a demand for approaches capable of tending to an ever-growing number of teaching staff who feel as if they are falling behind and incapable of keeping up and catching up with all that is being done regarding GenAI. This paper proposes a Strategic Framework for the ethical integration of Generative AI, developed through an exploratory initiative at the Viseu School of Education (ESEV), which has tackled these challenges and has worked towards the iterative and participative design of a strategic and scalable framework for the critical and its responsible integration of GenAI into faculty teaching practice. The work details the methodological architecture and operational indicators of this framework to ensure its scalability across higher education contexts. The multi-phase methodology adopted employs a robust mixed-methods approach, which includes a comprehensive assessment to diagnose faculty AI literacy, adoption patterns, and ethical practice and risk perceptions. Data collected informs the subsequent systematic inventory of high-utility tools and the establishment of a dedicated, interdisciplinary Study Group for continuous knowledge of co-creation. The project also includes the design and deployment of subject-specific, practical workshops focused on developing both the technical proficiency and critical pedagogical competencies required for effective integration of GenAI in teaching activi-

ties. It culminates in the delivery of evidence-based pedagogical recommendations and a comprehensive Faculty Manual for Ethical AI Integration. This project is not merely a context-specific case study but rather provides a transferable and potentially replicable framework for higher education institutions, aiming to enhance the quality of teaching and proactively cultivate an ethically grounded culture of innovation in the age of AI.

Keywords: Generative Artificial Intelligence, Higher Education, Pedagogical Innovation, Lecturer Training, AI Literacy.

1 Introduction

The rapid rise of Generative Artificial Intelligence (GenAI) is reshaping higher education, from assessment practices and curriculum design to faculty development, while creating asymmetric conditions in institutional readiness, access, and capacity [1] [2]. On the one hand, student adoption is accelerating; on the other, institutional responses and staff preparation remain uneven, deepening concerns about emerging “AI divides” in digital literacy and pedagogical innovation [3] [4] [5]. These tensions foreground a dual challenge: expanding pedagogical innovation, personalized learning, immediate feedback, and support for administrative and teaching tasks without compromising academic integrity and core educational values such as authorship, critical thinking, and fairness [1] [2] [3].

While the potential of GenAI to transform learning is widely acknowledged, recent literature highlights a critical gap in practical, scalable frameworks for faculty development that move beyond tool-focused training [6] [7]. Many initiatives remain confined to introductory workshops or ad hoc guidance, which help educators understand what GenAI can do, but provide little support for deciding what it should do within specific disciplinary and institutional contexts. As a result, lecturers frequently report a sense of being simultaneously pressured to adopt GenAI and underprepared to integrate it in ways that are pedagogically meaningful and ethically defensible [3] [4]. This gap between expectation and support reinforces uneven practices and may inadvertently widen internal disparities in AI literacy and confidence among staff.

At the same time, higher education institutions face growing external pressure to demonstrate that their responses to GenAI are aligned with broader regulatory and ethical frameworks. International guidance stresses the need to develop AI competencies among teachers, not only in technical terms but also in relation to ethics, governance, and human-centred design [6] [7]. The European policy landscape similarly emphasizes AI literacy as a strategic capability that must be cultivated across sectors, including education [8]. For institutions, this means that GenAI cannot be treated as a temporary disruption, but as a catalyst that requires rethinking professional development, assessment, and curriculum design in integrated ways.

It is against this background that the present study situates itself. The paper reports on an exploratory initiative implemented at the School of Education (ESEV) of the Polytechnic Institute of Viseu (IPV), which seeks to address the scarcity of evidence-based models for institutional capacity building in the age of AI. Rather than focusing

solely on compliance or on the demonstration of isolated use cases, the project aims to support lecturers in developing coherent, context-sensitive approaches to GenAI that are grounded in their own teaching realities. The core assumption is that sustainable innovation emerges when educators are positioned not merely as end-users of tools, but as co-designers of AI-enhanced learning environments.

To this end, the project adopts a Design-Based Research (DBR) approach that combines iterative design, implementation, and reflection with close collaboration between researchers and practitioners.

The goal is to co-create a support ecosystem that aligns pedagogical innovation with ethical stewardship, offering a replicable blueprint for integrating GenAI into teaching practices in ways that are both critically grounded and operationally feasible [7] [8]. Following this introduction, Section 2 synthesizes recent literature on GenAI in higher education from faculty and governance perspectives; Section 3 details the DBR methodology and its implementation phases; Section 4 discusses findings and expected outcomes; and Section 5 presents conclusions, limitations, and directions for future work.

2 Literature Review

2.1 Generative AI in Higher Education: Studies and Projects from Faculty Perspectives

Generative Artificial Intelligence (GenAI) systems use algorithms to analyze large datasets and generate original content in text, images, and audio formats [9]. This technology presents universities with both opportunities and challenges, prompting discussion among educators about its potential to transform teaching methods [10].

Research examining faculty perspectives reveals a complex landscape where educators recognize GenAI's transformative potential while grappling with significant implementation challenges [11]. Faculty members use these tools as virtual assistants that provide real-time personalized support and help develop educational resources tailored to student needs [11]. GenAI can enhance teaching through personalized learning and innovative pedagogy [12] [13], while improving accessibility through multimodal content [14].

However, academic integrity emerges as the most pressing concern. Faculty express significant worries about detecting and preventing academic dishonesty, particularly difficulties in identifying AI-generated plagiarism [2] [15] [16]. Beyond academic misconduct, educators identify broader ethical challenges including data privacy, security issues, potential bias from training datasets, and questions about whose expertise influences AI development [16] [17] [18].

Student over-reliance on AI tools represents another major concern, with faculty worried about impacts on soft skills and critical thinking capabilities [16] [19]. Semi-structured interviews reveal tensions between instructor and student preferences for AI engagement [20] [21] [22].

A critical systemic challenge is insufficient AI literacy and inadequate institutional resources among faculty, which significantly hinder effective adoption [23]. Successful

implementation requires comprehensive institutional transformation beyond mere technology adoption, including revisions to governance, assessment, and institutional values [24]. The APRU document [14] recommends comprehensive training programs combining technical skills with interdisciplinary discussions, along with institutional partnerships for sharing best practices. Faculty require ongoing professional development to understand GenAI's capabilities and limitations, especially regarding academic integrity [5] [25].

Institutional responses vary widely: some universities ban AI tools while others integrate them into active learning approaches [26]. Evidence suggests that equity-focused implementation can boost student motivation and enhance feedback quality, provided institutions balance innovation with accountability and preserve core academic principles [27] [28].

2.2 Building Technical and Pedagogical Competencies for AI-Enhanced Teaching

The successful pedagogical incorporation of GenAI hinges on faculty cultivating comprehensive digital competencies that encompass both technical knowledge of generative systems and awareness of associated ethical and regulatory boundaries. While many educators acknowledge the potential of GenAI to enhance classroom experiences and provide immediate responses, they also recognize gaps in their understanding of operational mechanics and optimal integration strategies that safeguard ethical standards and foster student originality [25] [27].

UNESCO [6] further emphasizes the potential of technology to catalyze collaborative creation between instructors and students, thereby fostering innovation and supporting learner autonomy. With appropriate preparation, educators can reimagine GenAI beyond simple content production, transforming it into an educational instrument that cultivates analytical reasoning, scholarly inquiry, and independent intellectual development.

The phenomenon of AI hallucinations, responses that appear plausible yet contain factual inaccuracies, underscores the importance of strengthening digital literacy among both instructors and learners when evaluating the credibility of machine-generated outputs [9] [28]. Such technical pedagogical expertise becomes crucial for educators to facilitate critical and accountable use of these resources, preventing uncritical reliance on automated content generation.

Recent evidence from faculty interviews reinforces that competence building must be treated as an institutional priority, not an individual burden. Management educators report feeling underprepared to guide students' responsible use of tools such as ChatGPT, citing steep learning curves, limited training opportunities, and uncertainty about how to integrate AI without compromising academic integrity or the mentor-student relationship [29].

These concerns extend to data privacy and compliance, where faculty call for more explicit institutional guidance on safe practices for student information and course materials, particularly in light of European data protection expectations [29].

Policy analyses also suggest that training should be designed as a structured, ongoing programme that develops both staff and students. Across a global sample of university policies and guidelines, standard measures include formal guidance on the ethical use of AI, authentic assessment design to reduce misuse, and capacity-building initiatives such as workshops, online modules, and dedicated support resources to foster AI literacy [30].

At the same time, gaps persist, especially around data privacy provisions and equitable access to tools, suggesting that faculty competencies must include practical knowledge of risk mitigation, disclosure norms, and inclusion-oriented teaching strategies [30].

From an organisational perspective, competence building is most effective when supported by units that can translate strategy into practice. Analyses of strategic communication at leading universities indicate that teaching and learning offices and development centres often serve as bridges between institutional direction and grassroots innovation, absorbing top-down priorities while enabling bottom-up experimentation and professional learning [31].

This suggests that faculty development can be strengthened through communities of practice, internal advisory groups, and accessible support structures that keep guidance current amid rapid tool evolution.

Finally, professional development efforts should anticipate uneven adoption across staff cohorts. Institutional guideline discourse highlights generational divides in digital literacy, where some faculty may require targeted support to adopt these tools confidently and ethically, reducing the risk that unequal readiness widens internal capability gaps [32].

2.3 Governance, Ethics, and Regulatory Frameworks for Responsible AI Integration

The integration of GenAI into higher education necessitates robust governance structures that strike a balance between innovation and ethical and regulatory responsibility. Recent comparative analyses have shown that universities across regions adopt varied approaches, often lacking coherent institutional frameworks to manage the adoption and risks of GenAI. [33] analyzed AI policies across 343 universities in five countries, revealing that more than half of the universities delegate AI use decisions to instructors, while fewer than one-fifth enforce strict prohibitions. Only 31% explicitly link AI policies to academic integrity, and almost half provide no clear guidance to staff or students. This fragmented landscape highlights the global lack of standardized governance models and underscores the need for adaptive institutional policies that are grounded in transparency, accountability, and inclusivity.

Similarly, [34] examined Swedish higher education institutions and found that only a minority had formal guidelines for GenAI. Most institutions offered limited web-based information rather than structured governance documentation. The study highlighted five critical policy dimensions: good academic practice, GenAI use and governance in education, information governance, ethical and social impact, and GenAI essentials. Swedish institutions emphasize transparency and responsibility; however, the

lack of national coordination creates inconsistencies in ethical standards and access to equity. The authors argue that universities must institutionalize AI literacy, integrate ethical training, and develop national directives that harmonize policy development across sectors.

Ethical governance of GenAI also hinges on managing risks related to data protection, algorithmic bias, and human oversight. [35] identified data privacy and algorithmic fairness as leading challenges in higher education. They emphasized compliance with the General Data Protection Regulation (GDPR) and the use of advanced encryption to protect student information. Their findings advocate institutional audits on algorithmic bias, ensuring diverse datasets and transparent decision-making to prevent discrimination. Furthermore, the authors propose clear institutional policies defining acceptable AI use to preserve academic integrity and authenticity in student work.

At a broader level, governance frameworks should connect institutional policy to social responsibility and sustainable innovation. [1] observed that effective GenAI governance depends on embedding pedagogical theory, such as Laurillard's Conversational Framework and the SAMR model, into policy design. These models foster reflective and interactive learning while ensuring the ethical and inclusive use of technology. Embedding ethical awareness in governance allows higher education institutions to navigate tensions between innovation, integrity, and equity.

In parallel, [36] described a tri-phased approach to integrating GenAI in higher education that combines literature review, stakeholder consultation, and qualitative data analysis. Their research demonstrates that a transparent and participatory governance process enhances policy acceptance among faculty and mitigates ethical concerns related to plagiarism and loss of interpersonal skills. Governance structures should therefore promote stakeholder participation, including faculty, students, and administrative units, to ensure policies remain contextually relevant and ethically grounded.

Overall, the literature converges on three governance imperatives: (1) formal institutional policies aligned with global and regional regulatory frameworks such as the GDPR and the forthcoming EU AI Act, (2) explicit ethical standards addressing fairness, privacy, and integrity, and (3) capacity-building mechanisms that enhance faculty AI literacy and ethical reasoning. Effective governance of GenAI in higher education must thus move beyond reactive regulation toward a proactive culture of ethical innovation, ensuring that AI serves as a pedagogical ally while upholding the principles of transparency, responsibility, and human agency.

3 Methodology

This project adopts a Design-Based Research (DBR) methodology for tackling both the goal of solving a pragmatic educational problem, the lack of structured integration of Generative AI (GenAI) in teacher training, and the generation of theoretical insights into how these technologies influence pedagogical innovation. These may be considered the challenges established at the beginning of the project, and that will culminate in multiple outcomes that may go far beyond the provision of guidelines and training

activities suitable for teachers. DBR is also considered inherently iterative and interventionist, so it is considered ideal for the complex environment of higher education teaching and learning activities, where technological, professional, and institutional variables are continuously combined and rethought.

The iterative nature of this process demands solutions like the ones provided by employing the Framework for Innovation Design (the Double Diamond) [37], which segments the research into four distinct phases: Discover, Define, Develop, and Deliver, begins with a challenge, ends with an outcome and may have more than one iterative moment during the 4 phases before mentioned. This methodological workflow ensures that the technical requirements of GenAI systems, such as API constraints, data privacy protocols, and prompt engineering architectures, are systematically integrated into the pedagogical design.

This 4-phase structure is believed to ensure a rigorous progression between macro activities, such as mapping teachers need to establish sustainable communities of practice and utilizing a mixed-methods approach for data collection. The macro activities throughout the project are illustrated in Fig. 1.



Fig. 1. Macro activities thought-out for the project.

In the first stage, known as the Discover stage, the project team will focus on understanding the context and completing the "Map of Teacher Perceptions." This mapping initiative will outline what teachers already know and what they need to learn about AI. To do these quantitative surveys, they will be applied online and distributed among the entire teaching staff at ESEV, who will hopefully check and share their perspective regarding their own current skills and AI literacy. At the same time, semi-structured interviews with a smaller group of teachers from different scientific fields will be conducted. These interviews are important to understand the participants' personal feelings, fears, and ideas about using AI in their teaching activities and higher education in general. This stage will also help them to become self-aware of the gaps in their own professional training.

In the Define stage, the information collected in the first stage will be used to make decisions. One of its goals is to create an "Inventory of GenAI Tools", according to their perceived educational values. To do this, the project includes a comparative analysis of a list of different AI applications. A focus group with teaching staff will discuss the results of this comparative analysis and help choose a set of tools considered useful

for the project training activities. The resulting comparative matrix and the remaining inherent activities will help define the guidelines to be further defined in the project.

The third stage is Develop, and it is probably the most creative part of the project. The "ESEV Online Manual" and the "Pilot Lesson Plans" will both be produced at this stage. The outcomes of this stage are intended to be used as materials that may help teachers use AI in critical and innovative ways when compared to their current practices. The manual will be reviewed and commented on by a group of specialists selected according to their work already conducted in the field. The lesson plans will test them in a real class environment, and during these tests, the activities will be directly observed in the classes to determine what works and what does not work according to the lesson plan's expected results. This feedback will aid the revision and improvement of the materials designed before their final implementation.

The four stages end with the Deliver moment that focuses on the result and the "Establishment of the Study Circle." The project's impact will be measured with the use of an online questionnaire that will touch base in three areas: student learning, teacher skills, and institutional strategy, as illustrated in Fig. 2.



Fig. 2. Three main areas of the project.

Each of these levels will be evaluated according to specific indicators and Performance Indicators exemplified in table 1.

Table 1. Examples of evaluation indicators.

Level	Dimensions	Key Performance Indicators (KPIs)
Institutional	Strategic Alignment	Number of departmental guidelines updated; Frequency of AI-related mentions in Quality Assurance reports.
Professional	AI Literacy	Change in pre- and post-training technical proficiency scores; Number of AI-integrated lesson plans peer-reviewed.
Social & Student	Pedagogical Impact	Percentage of students reporting critical thinking when engaging with AI outputs; Reduction in academic integrity incidents related to AI misuse.

To do so, teachers will be tested before and after the training, through an online questionnaire, as to their AI literacy, opinion about their teaching and learning activities and institutional strategy, and the team will try to assess the project's impact in these fields. Teachers will also be asked to share some personal input throughout their journey in a form with open questions about their endeavors. All the data collected and analyzed is expected to contribute to solving the challenge and delivering the outcome aimed for. It is hoped that, in the end, the project may help create at ESEV a culture and concern with the improvement and innovation in terms of AI literacy in the teaching staff and how it propagates to teaching and learning activities and the institution's overall strategy in relation to AI.

4 Expected Results and Impact

The project is expected to generate a coherent set of results that operate at the levels of classroom practice, professional culture, and institutional strategy within ESEV. At a practical level, the initiative aims to consolidate resources that make the integration of GenAI into teaching activities more intentional, transparent, and sustainable. Rather than leaving each lecturer to navigate GenAI independently, the project will provide shared reference points that reduce fragmentation and support more consistent decision-making across programmes and disciplines.

One key expected result is a clearer articulation of how GenAI can be aligned with existing curricular goals and assessment practices. By working with lecturers from different subject areas, the project intends to identify patterns of use that genuinely add pedagogical value, as opposed to simply replicating existing tasks with new tools. This is anticipated to result in a portfolio of example activities and assessment strategies that demonstrate how GenAI can support inquiry, formative feedback, and discipline-specific skills, while still preserving space for student authorship and critical thinking. The emphasis is not only on what tools are available, but on what kinds of learning they enable or constrain in concrete classroom situations.

A second cluster of expected results concerns the professional development of lecturers. Through their involvement in iterative cycles of design, trial, and reflection, teachers are expected to move from isolated experimentation or avoidance towards a more confident and principled engagement with GenAI. This should be visible in several ways: in the language they use to discuss GenAI with colleagues and students; in the explicitness with which they frame expectations around its use; and in their ability to justify, adapt, or suspend GenAI-based activities in response to pedagogical or ethical concerns. The project thus aims to foster a shift from viewing GenAI primarily as a technical challenge to understanding it as a matter of pedagogical design and professional judgement.

At the institutional level, the project is expected to contribute to a more strategic and anticipatory stance towards GenAI. Drawing on the evidence gathered throughout the

different phases, the team will be able to outline principles and processes that can inform future policy discussions, accreditation processes, and quality assurance mechanisms. This may include, for example, procedures for periodically updating guidance as tools evolve, or for involving students and external stakeholders in conversations about acceptable AI use. In this sense, the project's results are expected to outlive the specific tools currently in use, providing a framework that remains relevant even as platforms become obsolete.

For students, the anticipated impact is to encounter learning environments where GenAI is neither uncritically celebrated nor simply banned, but integrated in ways that are explicit, purposeful, and open to discussion. As lecturers gain confidence and clarity, students should benefit from tasks that require them to engage with GenAI critically, questioning outputs, comparing perspectives, and reflecting on their own role in the knowledge-production process. Over time, this is expected to support the development of more robust digital and AI literacies, better preparing students for contexts beyond higher education where such competencies will be increasingly necessary.

Taken together, these results and impacts aim to demonstrate that it is possible to integrate GenAI into higher education in a way that supports innovation without eroding the core values of academic work. The project seeks not only to respond to current pressures but to lay the groundwork for a more reflective, participatory, and ethically grounded approach to future technological change.

Taken together, these results and impacts aim to demonstrate that it is possible to integrate GenAI into higher education in a way that supports innovation without eroding the core values of academic work. The project seeks not only to respond to current pressures, but to lay the groundwork for a more reflective, participatory, and ethically grounded approach to future technological change.

5 Conclusions/Final considerations and Future Work

The integration of Generative Artificial Intelligence into higher education is neither a purely technological nor a purely pedagogical undertaking; it is a multidimensional process that demands ethical vigilance, institutional coherence, and sustained professional development. This study demonstrates that successful adoption should hinge itself on a strategic approach that prioritizes human agency over technological determinism. Future results, namely the proposal of a strategic framework capable of translating comprehensible ethical guidelines into actionable pedagogical indicators, are expected to contribute to current practices not only within ESEV, but also in partner higher education institutions still finding their way through all the noise and hype gravitating around GenAI related issues and practices. By providing a structured progression, from diagnostic mapping to the establishment of Study Circles, ongoing work will offer the basis for replicable practices regarding institutional AI readiness. By embedding GenAI within a framework of critical reflection and collaborative learning, the initiative at ESEV-IPV illustrates how institutions can transform disruption into opportunity.

In this context, the integration of GenAI in higher education combines high pedagogical potential [12] [13] with ethical and assessment challenges that raise questions

about the role of the teacher [14] [38]. This project addresses these challenges through a proactive and structured approach, recognizing that technology alone does not ensure innovation. Successful integration depends on continuous teacher training, institutional regulatory adaptation, and a culture of integrity and critical thinking [14]. The project aligns with international initiatives promoting AI literacy and ethical GenAI use in teaching [39] [40] [41] [42]. By combining diagnosis, training, and applied research, the project provides a contextualized and transferable model, aligned with the EU Artificial Intelligence Act [8] and supporting more equitable digital transformation [6] [14].

Several key insights emerge from this work. First, faculty development is indispensable. Educators must acquire not only operational proficiency but also the capacity to interrogate the epistemological and ethical implications of GenAI. Without this dual competence, the risk of superficial adoption and compromised academic integrity remains high. Second, institutional alignment is critical. Fragmented policies and inconsistent practices undermine trust and exacerbate disparities in digital readiness. Clear guidelines, informed by international standards and adapted to local contexts, are essential to ensure coherence and accountability. Third, ethical considerations must permeate every stage of integration. Issues such as data privacy, algorithmic bias, and intellectual authorship are not peripheral concerns; they are central to safeguarding the values upon which higher education rests.

The project's outcomes, such as diagnostic assessments, curated tool inventories, interdisciplinary study circles, and evidence-based manuals, constitute a robust foundation for institutional transformation. Yet, these achievements should be understood as provisional rather than definitive. The rapid evolution of GenAI technologies and regulatory frameworks necessitates continuous adaptation. Moreover, the sustainability of such initiatives depends on resource allocation, leadership commitment, and the cultivation of a culture that values innovation without sacrificing integrity.

In sum, this project affirms that the integration of GenAI can enrich teaching and learning when guided by principles of responsibility, inclusivity, and critical engagement. It offers a replicable model for institutions seeking to navigate the dual challenge of technological advancement and ethical stewardship, contributing to the broader discourse on the future of higher education in the age of artificial intelligence.

Yet, despite these contributions, important questions remain open, pointing toward avenues for further inquiry.

Future research should extend beyond the institutional boundaries of this study to explore comparative and collaborative approaches. Establishing inter-university networks and communities of practice can accelerate knowledge exchange and harmonize standards across diverse educational contexts. Longitudinal studies are necessary to assess the impact of GenAI integration on student learning outcomes, faculty workload, and institutional reputation, providing empirical evidence to inform policy and practice.

Another priority is the development of adaptive assessment models that incorporate GenAI without compromising academic integrity. This includes designing evaluation frameworks that recognize the legitimate use of AI tools while preserving originality and critical thinking. Additionally, future initiatives should investigate the intersection

of GenAI with equity and inclusion, ensuring that technological innovation does not exacerbate existing disparities in access and opportunity.

Emerging technologies—such as multimodal AI systems and personalized learning platforms—will further complicate the ethical and pedagogical landscape. Anticipating these developments requires proactive scenario planning and continuous professional development. Ultimately, future work must embrace flexibility, interdisciplinarity, and evidence-based adaptation, positioning higher education as a leader in shaping an ethically grounded digital future.

References

1. Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Hassan, G.M.: Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence* **8**, 100407 (2025)
2. Cotton, D.R.E., Cotton, P.A., Shipway, J.R.: Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* **61**(2), 228–239 (2024)
3. Farrokhnia, M., Banihashem, S.K., Noroozi, O., Wals, A.: A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International* **61**(3), 460–474 (2024)
4. Perkins, M.: Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice* **20**(2), Article 07 (2023)
5. Rudolph, J., Tan, S., Tan, S.: ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching* **6**(1), 342–363 (2023)
6. UNESCO: AI competency framework for teachers. UNESCO (2024)
7. Miao, F., Holmes, W.: Guidance for generative AI in education and research. UNESCO (2023)
8. European Union: Artificial Intelligence Act – Article 4: AI literacy. <https://artificialintelligenceact.eu/article/4/>, last accessed 2025/02/01
9. OpenAI: GPT-4 Technical Report. arXiv:2303.08774 (2023). <https://arxiv.org/pdf/2303.08774>
10. Rudolph, J., Tan, S., Tan, S.: War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning & Teaching* **6**(1), 364–376 (2023)
11. Ruiz-Rojas, L.I., Acosta-Vargas, P., De-Moreta-Llovet, J., Gonzalez-Rodriguez, M.: Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability* **15**, 11524 (2023)
12. Kasneci, E., Sessler, M., Schramm, K., Mayer, S.: ChatGPT for good? On opportunities and challenges of large language models for education. *Frontiers in Artificial Intelligence* **6** (2023)
13. Baidoo-Anu, D., Owusu Ansah, L.: Education in the era of generative artificial intelligence: Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI* **7**(1), 52–62 (2023)

14. APRU: Generative AI in higher education: Current practices and ways forward. White Paper (2025). https://www.apru.org/wp-content/uploads/2025/01/APRU-Generative-AI-in-Higher-Education-Whitepaper_Jan-2025.pdf
15. AL-Smadi, M.: ChatGPT and Beyond: The Generative AI Revolution in Education. arXiv:2311.15198 (2023)
16. Jomezai, N.A., Baloch, F.A., Jaffar, M., Khilji, G.: Generative AI in higher education: University faculty perspectives on opportunities and challenges. *Turkish Online Journal of Distance Education* **26**(3), 128–143 (2025)
17. Lindsay, E.D., Zhang, M., Johri, A., Bjerva, J.: The responsible development of automated student feedback with generative AI. arXiv:2308.15334 (2023)
18. Aad, S., Hardey, M.: Generative AI: Hopes, controversies and the future of faculty roles in education. *Quality Assurance in Education* **33**(2), 267–282 (2025)
19. Vrågård, J., Brorsson, F., Aghaee, N.: Generative AI in higher education: Educators' perspectives on academic learning and integrity. In: *Proceedings of the 23rd European Conference on e-Learning (ECEL 2024)*, pp. 406–414. Academic Conferences International (2024)
20. Zastudil, C., Rogalska, M., Kapp, C., Vaughn, J.L., MacNeil, S.: Generative AI in computing education: Perspectives of students and instructors. arXiv (2023)
21. Becker, B.A., Denny, P., Finnie Ansley, J., Luxton Reilly, A., Prather, J., Santos, E.A.: Programming is hard, or at least it used to be: Educational opportunities and challenges of AI code generation. In: *Proceedings of the 54th ACM Technical Symposium on Computer Science Education*, pp. 500–506. ACM (2023)
22. Prather, J., et al.: Transformed by transformers: Navigating the AI coding revolution for computing education. In: *ITiCSE 2023 Conference on Innovation and Technology in Computer Science Education*, vol. 2, pp. 561–562. ACM (2023)
23. Bastý, R., Kropczynski, J., Halse, S.: Exploring higher education faculty insights on generative AI in creative courses. *Journal of Information Technology Education: Research* **24**, 18 (2025)
24. EDUCAUSE: 2024 EDUCAUSE Horizon Report: Teaching and Learning Edition. EDUCAUSE, Boulder (2024). <https://www.educause.edu/horizon-report-teaching-and-learning-2024>
25. Kamalov, F., Santandreu Calonge, D., Gurrib, I.: New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability* **15**(16), 12451 (2023)
26. Kasneci, E., Sessler, K., Schramm, K., Mayer, S., et al.: ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* **103**, 102274 (2023)
27. Kurtz, G., et al.: Strategies for integrating generative AI into higher education: Navigating challenges and leveraging opportunities. *Education Sciences* **14**(5), 503 (2024)
28. Lo, C.K.: What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences* **13**(4), 410 (2023)
29. Shahzad, F., Avunduk, Z.B., Ullah, Z., Arslan, A.: Understanding management educators' perspectives on AI adoption: A qualitative study. *International Journal of Management Education* **23**(3), 101274 (2025)
30. Jin, Y., Yan, L., Echeverria, V., Gašević, D., Martinez-Maldonado, R.: Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence* **8**, 100348 (2024)
31. Gering, Z., Feher, K., Harmat, V., Tamassy, R.: Strategic organisational responses to generative AI-driven digital transformation in higher education institutions. *International Journal of Organizational Analysis* **33**(12), 132–152 (2025)

32. Alzubaidi, K.: The role of Generative AI in higher education: Institutional guidelines, generational gaps, and the GROK 4 challenge. *Arab World English Journal* **11**, 1–4 (2025)
33. Parker, L., et al.: Comparative analysis of artificial intelligence policies in universities across five countries. *Discover Computing* **28**, 267 (2025)
34. Erhardt, C., et al.: From policy to practice: The regulation and implementation of generative AI in Swedish higher education institutes. *International Journal for Educational Integrity* **21**(21) (2025)
35. García-Lopez, I.M., Gonzalez, C.S., Ramírez-Montoya, M.S., Molina-Espinosa, J.M.: Challenges of implementing ChatGPT in education: Systematic literature review. *International Journal of Educational Research Open* **8**, 100401 (2025)
36. Kumar, S., Rao, P., Singhania, S., Verma, S., Kheterpal, M.: Will artificial intelligence drive the advancements in higher education? A tri-phased exploration. *Technological Forecasting and Social Change* **201**, 123258 (2024)
37. Design Council: Framework for Innovation. Design Council Website. Available from: <https://www.designcouncil.org.uk/our-resources/framework-for-innovation/> (2025), last accessed 2025/11/23
38. Hasanein, A.M., Sobaih, A.E.E.: Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health, Psychology and Education* **13**(11), 2599–2614 (2023)
39. Baytas, C., Ruediger, D.: Making AI generative for higher education: Adoption and challenges among instructors and researchers. *Ithaka S+R* (2025). <https://sr.ithaka.org/publications/making-ai-generative-for-higher-education/>
40. Vanderbilt University: Teaching with Generative AI (2025). <https://www.vanderbilt.edu/generative-ai/teaching/>, last accessed 2025/02/01
41. University of Texas at Austin: Faculty Guide to Getting Started with Generative AI. Office of Academic Technology (2023). <https://content.grammarly.com/view/193894986/18/>, last accessed 2025/02/01
42. Bowskill, N., Hall, D., Harrogate, M., Eziefuna, E., Marler, B.: Staff development for generative artificial intelligence and collaborative learning using Iterationism as a theoretical framework. *Journal of Learning Development in Higher Education* **33** (2025)

Aspetos audiovisuais na construção de vídeos de exercício físico para pessoas mais velhas com dor crónica

Ana Teresa Bessa¹, Telmo Silva¹ [0000-0001-9383-7659],
and Anabela G. Silva² [0000- 0002-4386-5851]

¹ DigiMedia, Department of Communication and Art,
University of Aveiro, Aveiro, Portugal

² RISE-Health, School of Health Sciences, University of Aveiro, Aveiro, Portugal
{anateresabessa, tsilva, asilva}@ua.pt

Resumo. Numa sociedade em que o digital é, cada vez mais, uma presença constante, é importante compreender como as diferentes representações visuais de um mesmo conteúdo (por exemplo: uma pessoa real ou um avatar digital) podem influenciar a compreensão e a motivação das pessoas mais velhas na prática de exercício físico. A tecnologia na área da saúde é, sobretudo, desenvolvida para utilizadores mais jovens e familiarizados com o meio digital, deixando, comumente, de parte o público sénior. Este estudo tem como objetivo avaliar a perceção das pessoas idosas com dor crónica relativamente à forma de apresentação dos exercícios físicos prescritos pelos seus fisioterapeutas a partir de uma plataforma. A metodologia utilizada tem como base numa abordagem qualitativa e participativa, estruturada em dois *workshops* com pessoas mais velhas. Em cada *workshop* serão apresentados vídeos com a demonstração de três formas de apresentação dos exercícios, uma protagonizada por um instrutor humano, uma por um avatar digital e outra criada por inteligência artificial mantendo-se o conteúdo e instruções. Através da observação direta e da realização de um *focus group*, serão recolhidas as perceções, as preferências e as reações dos participantes. Estes resultados vão contribuir para o desenvolvimento dos conteúdos audiovisuais, reforçando o papel do utilizador no desenvolvimento dos meios audiovisuais na comunicação em saúde.

Palavras-chave: Pessoas mais velhas; Dor crónica; Saúde digital; Exercício físico; Avatar digital; Inteligência Artificial.

1 Introdução

O envelhecimento populacional tem levado a um aumento da dor crónica, que afeta a funcionalidade, a autonomia e a qualidade de vida das pessoas mais velhas [1]. O aumento da prevalência da dor crónica entre adultos mais velhos tem conduzido a uma maior necessidade de acompanhamento clínico e de intervenções regulares, contribuindo para uma pressão acrescida sobre os serviços de saúde o que tem motivado a procura de soluções complementares de reabilitação [2][3]. Estudos mostram que condições como a dor lombar crónica representam um dos principais motivos para a

procura de cuidados de reabilitação e que a perda de funcionalidade associada a esta condição aumenta a dependência de apoio formal e informal [2][1].

Neste contexto, abordagens digitais, como a reabilitação remota, aparecem como alternativas para melhorar o acesso ao exercício e reduzir a carga sobre os serviços presenciais [3]. Contudo, apesar do potencial destas soluções, permanece por compreender de que forma as pessoas mais velhas percecionam as diferentes representações visuais utilizadas nos vídeos de exercício físico, bem como os elementos audiovisuais que compõem estas experiências.

Estudos recentes sugerem que características como o realismo, a presença social, a credibilidade e a identificação com a figura representativa influenciam a motivação e a adesão das pessoas mais velhas em contextos digitais de exercício físico [4][5]. Do mesmo modo, a utilização de avatares digitais apresenta vantagens em termos de personalização e envolvimento, mas também limitações relacionadas com a naturalidade do movimento, a confiança e a clareza motora [6][7].

Assim, é essencial investigar quais os elementos audiovisuais, figura, instruções, que melhor apoiam a compreensão, a segurança e a motivação das pessoas mais velhas com dor crónica na prática de exercício físico.

2 Enquadramento Teórico

A literatura demonstra que soluções digitais de exercício físico e reabilitação remota têm mostrado resultados positivos na melhoria da função física, na adesão ao exercício e na perceção de apoio emocional em pessoas mais velhas [2][3].

Vários autores referem que pessoas mais velhas tendem a valorizar a clareza, a simplicidade, a segurança e a credibilidade da figura representativa em contextos digitais de saúde [4][8][5][9]. Diversos estudos analisam a aceitação e o impacto de avatares digitais e robots como mediadores de exercício físico, revelando benefícios, mas também limitações, sobretudo ao nível da naturalidade do movimento, identificação emocional e confiança [4][8][5]. Estudos sobre avatares e figuras digitais sugerem que aspetos como antropomorfismo - consiste na atribuição de características humanas, como a aparência, a expressão, a voz, o comportamento intencional, a entidades artificiais ou digitais (por exemplo, avatares digitais ou robots) - o realismo e a naturalidade do movimento, podem influenciar a experiência e a motivação em contextos digitais de exercício físico [6][7][10]. No entanto, ainda são escassos os estudos que analisam as perceções de pessoas mais velhas com dor crónica perante diferentes figuras e estilos de conteúdo audiovisual, o que justifica a pertinência deste estudo.

Em contextos de saúde e reabilitação digital, níveis mais elevados de antropomorfismo tendem a influenciar positivamente a perceção de presença social, a confiança, a aceitação da tecnologia e o envolvimento dos utilizadores mais velhos, embora possam também gerar expectativas irreais quando o comportamento não corresponde ao realismo visual apresentado [4][11][12].

Apesar do crescente número de estudos sobre tecnologia e envelhecimento, permanece escassa a investigação que compara diretamente diferentes figuras

demonstrativas (humana, avatar digital e figura criada por IA) em vídeos de exercício físico especialmente direcionados a pessoas mais velhas com dor crónica. Do mesmo modo, ainda são limitadas as análises sobre a forma como elementos audiovisuais e instruções influenciam não apenas a compreensão, mas também a confiança e intenção de continuidade da prática de exercício físico.

3 Metodologia de Investigação

A metodologia deste estudo assenta numa abordagem qualitativa e participativa, estruturada de forma sequencial, articulando momentos de conceção, testes, produção e avaliação dos conteúdos audiovisuais. O processo metodológico desenvolve-se em fases interligadas, permitindo uma compreensão progressiva das preferências das pessoas mais velhas relativamente às figuras de representação e às instruções presentes nos vídeos de apoio à prática de exercício físico.

A fase inicial corresponde à contextualização da ideia, à definição do enquadramento conceptual e metodológico do estudo, incluindo a revisão da literatura, a identificação das principais lacunas de investigação e a definição dos critérios de inclusão dos participantes, na qual é definido o enquadramento teórico e delineadas as orientações gerais do estudo, bem como os critérios de inclusão dos participantes. Nesta etapa procede-se à produção dos diferentes formatos da figura de representação dos exercícios, criando três versões audiovisuais do mesmo exercício físico, uma com uma pessoa humana real, outra com um avatar digital e uma terceira com uma figura humana criada por Inteligência Artificial. Estas versões constituem o estímulo central para o teste com o público-alvo, pessoas com 65 ou mais anos com dor crónica.

Seguidamente, realiza-se os testes com os participantes, através de um workshop em formato de focus group (os resultados deste workshop serão descritos na secção 5 deste documento). Durante o workshop, as três versões dos vídeos serão apresentadas sequencialmente, seguindo-se momentos orientados de discussão e reflexão onde os participantes são convidados a partilhar as suas perceções, preferências e interpretações relativamente aos conteúdos apresentados. Paralelamente, é efetuada recolha de dados por observação direta, registando reações comportamentais, expressões e dificuldades manifestadas pelos participantes. Este formato permite promover a discussão coletiva e explorar de forma aprofundada as experiências e expectativas dos participantes.

Complementarmente, é aplicado um questionário de resposta simples, permitindo recolher perceções individuais relativamente a diferentes dimensões da experiência de visualização, tais como a clareza das instruções, conforto, identificação com a figura de representação, a preferência entre os diferentes formatos apresentados, a compreensão dos movimentos e a confiança transmitida pela figura demonstrativa.

Os dados recolhidos são posteriormente sujeitos a análise de conteúdo, procurando identificar padrões de preferência, dificuldades percebidas e aspetos valorizados pelos participantes. Estes resultados orientam um processo de revisão e avaliação dos conteúdos audiovisuais produzidos.

A fase seguinte corresponde ao processo de pré-produção, produção e pós-produção dos vídeos, no qual são incorporadas as recomendações resultantes da etapa anterior,

assegurando a maior adaptação das escolhas sobre os aspetos visuais, sonoros e das instruções relativamente às necessidades do público-alvo. Esta fase inclui ainda novas revisões internas, assegurando rigor técnico, acessibilidade e coerência.

Por fim, os conteúdos finalizados são integrados na plataforma digital e procede-se à realização de testes de usabilidade com o mesmo tipo de participantes. Nesta etapa pretende-se avaliar não apenas a compreensão das instruções e a clareza dos movimentos, mas também a capacidade de utilização da plataforma, a execução dos exercícios e o nível de conforto e confiança demonstrados pelos participantes. Novamente, a recolha de dados articula observação direta e instrumentos de resposta simples, permitindo uma análise qualitativa. Os dados obtidos são analisados e constituem a base para a reflexão final sobre a eficácia dos conteúdos produzidos.

Esta metodologia procura articular a investigação participativa e a produção de conteúdos audiovisuais, garantindo que os conteúdos desenvolvidos respondem às necessidades reais dos participantes, as pessoas mais velhas com dor crónica, e contribuem para melhorar a experiência da prática de exercício físico em contexto digital.

4 Resultados Esperados

Espera-se que o estudo permita identificar, de forma clara, quais os elementos audiovisuais e características da figura de representação humana nos vídeos, que mais contribuem para a compreensão, a segurança, a motivação e a adesão das pessoas mais velhas com dor crónica à prática de exercício físico. Sabe-se, previamente, que existem diferenças na perceção entre as três modalidades de representação, pessoa real, avatar digital e figura criada por inteligência artificial, nomeadamente ao nível da confiança transmitida, da credibilidade das instruções, da naturalidade do movimento e do conforto emocional.

Prevê-se que a figura humana real possa ser percecionada como mais próxima, mais empática e confiável, favorecendo a identificação e a sensação de segurança na execução do movimento do exercício físico. Espera-se recolher dados relevantes sobre as preferências relacionadas com o enquadramento, ritmo, som, texto e características tipográficas, identificando boas práticas na construção de vídeos acessíveis e ajustados às necessidades deste público-alvo.

Do ponto de vista metodológico, os resultados deverão contribuir para orientar a produção dos conteúdos audiovisuais em contexto de saúde, especialmente direcionados a pessoas mais velhas com dor crónica.

5 Resultados Preliminares do Primeiro *Workshop*

A primeira fase da investigação consistiu na realização de um workshop com sete pessoas mais velhas (de idade igual ou superior a 65 anos) com dor crónica, no qual foram exploradas as perceções e as preferências relativamente aos diferentes formatos de apresentação da figura de representação dos vídeos de exercícios físicos. Este

momento permitiu recolher evidências qualitativas relevantes sobre a usabilidade, clareza das instruções e adequação das figuras representativas.

Os *workshops* têm uma duração aproximada de 60 a 90 minutos, decorrendo em ambiente controlado de forma a garantir condições adequadas de visualização e audição dos conteúdos apresentados.

No início da sessão foi realizada uma breve apresentação do estudo, explicando aos participantes os objetivos da investigação, a estrutura da atividade e o funcionamento da sessão. Seguidamente foram apresentados os três vídeos demonstrativos do exercício físico, correspondentes às diferentes formas de representação visual utilizadas na investigação.

Os vídeos foram apresentados de forma sequencial, mantendo constantes todos os elementos audiovisuais relevantes, tais como duração do exercício, enquadramento da câmara e instruções, variando apenas a figura demonstrativa (instrutor humano, avatar digital ou figura criada por inteligência artificial). Esta estratégia permitiu garantir a comparação entre as diferentes versões apresentadas.

Após a visualização de cada vídeo, os participantes foram convidados a partilhar as suas impressões iniciais relativamente à clareza das instruções, à naturalidade do movimento, à confiança transmitida pela figura demonstrativa e à facilidade de compreensão do exercício apresentado.

Posteriormente foi conduzida uma discussão em formato de *focus group*, moderada pelo investigador, na qual os participantes puderam discutir as suas perceções de forma mais aprofundada, comparando as diferentes versões apresentadas e justificando as suas preferências.

Durante toda a sessão foram registadas observações relativas às reações dos participantes, incluindo comentários espontâneos, expressões faciais e eventuais dificuldades de compreensão ou interpretação dos conteúdos apresentados.

Os resultados revelaram uma preferência pela figura humana real como forma de representação dos vídeos de exercício. Os participantes associaram esta opção a maior naturalidade do movimento, confiança, identificação e sensação de segurança na execução. Em contraste, a versão de avatar digital foi unanimemente a menos apreciada, sendo descrita como artificial, pouco expressiva e com menor componente humana. Também a figura criada por Inteligência Artificial gerou algumas controvérsias, sobretudo no que respeita à expressão facial, naturalidade e postura.

Foram ainda identificados critérios estéticos e técnicos valorizados pelos participantes como cenário simples e não distrativo (com preferência por fundo branco), postura corporal correta da figura, ritmo adequado e uma aparência visual considerada “agradável” e “motivadora”. A roupa desportiva foi referida como preferível, associada à credibilidade do exercício.

Relativamente às instruções, os participantes destacaram a necessidade de clareza, simplicidade e objetividade, defendendo mensagens curtas, diretas e visuais, complementadas por narração com voz humana. Demonstraram preferência por informação visível no ecrã acompanhada por uma contagem verbal, defendendo também a importância de feedback positivo (verbal ou visual) no final do exercício. Foi, ainda, valorizada a existência de orientações prévias, tais como calçado adequado,

hidratação e uso de material (por exemplo, pesos), bem como existir flexibilidade no envio de lembretes relacionados com a hora da prática do exercício.

Do ponto de vista motivacional, foi enfatizada a importância da componente humana, mesmo em contexto digital, reforçando a ideia de que as soluções tecnológicas são percecionadas como complemento e não substituição da interação presencial de sessões de fisioterapia.

Com base nestes resultados preliminares, a próxima etapa do estudo consistirá na produção dos vídeos de apoio ao exercício físico, integrando as recomendações dos participantes relativamente à figura representativa, elementos audiovisuais, clareza das instruções, cenário visual e ritmo de execução.

Num momento seguinte, será realizada a fase de avaliação em contexto de utilização, recorrendo a testes de usabilidade com pessoas mais velhas que irão interagir com os conteúdos audiovisuais integrados numa plataforma digital. Esta fase permitirá analisar a compreensão do exercício, a capacidade de execução do exercício físico até ao fim e a experiência de usabilidade da plataforma digital.

Pretende-se que estes resultados permitam aprofundar o conhecimento sobre a adaptação de conteúdos audiovisuais para pessoas mais velhas com dor crónica, contribuindo para a criação de soluções digitais mais inclusivas, eficazes e centradas no utilizador.

6 Recolha e Análise de Dados

A recolha de dados baseou-se na utilização de três fases, permitindo obter uma compreensão mais abrangente da experiência dos participantes.

A primeira fase correspondeu à observação direta, realizada durante a visualização dos vídeos e ao longo da discussão em grupo. Esta observação permitiu registar reações comportamentais, comentários espontâneos e indicadores não verbais relevantes para a interpretação da experiência dos participantes.

A segunda consistiu na realização de um *focus group*, no qual os participantes foram convidados a discutir as suas perceções relativamente aos diferentes formatos de representação apresentados nos vídeos. Este método permitiu explorar em maior profundidade as opiniões dos participantes e compreender as razões subjacentes às suas preferências.

A terceira correspondeu à aplicação de um questionário estruturado, composto por perguntas de resposta fechada e escalas do tipo Likert. O questionário procurou avaliar diferentes dimensões da experiência de visualização, nomeadamente: clareza das instruções apresentadas nos vídeos; facilidade de compreensão dos movimentos; nível de confiança transmitido pela figura demonstrativa; identificação com a personagem apresentada; preferência entre os diferentes formatos de representação. Para além das questões fechadas, o questionário incluiu também campos de resposta aberta, permitindo aos participantes expressar comentários adicionais ou justificar as suas escolhas.

Os dados recolhidos ao longo da investigação foram analisados através de uma abordagem qualitativa, permitindo identificar padrões de significado nas perceções e experiências dos participantes.

O processo de análise seguiu várias etapas. Numa primeira fase foi realizada uma leitura aprofundada dos dados recolhidos, incluindo as notas de observação e os registos das intervenções realizadas durante o *focus group*. Esta etapa permitiu uma familiarização inicial com os conteúdos e a identificação de temas recorrentes nas perceções dos participantes.

Posteriormente foram identificados padrões de preferência entre as diferentes formas de representação dos exercícios físicos. Por fim, foi realizada uma interpretação dos resultados, relacionando os temas emergentes com a literatura existente sobre antropomorfismo, presença social e comunicação audiovisual em contextos de saúde digital.

Esta abordagem permitiu compreender não apenas as preferências manifestadas pelos participantes, mas também os fatores que influenciam a forma como as diferentes representações audiovisuais são percebidas em contextos práticos de exercício físico para pessoas mais velhas.

A presente investigação foi conduzida respeitando os princípios éticos aplicáveis à investigação com participantes humanos. Todos os participantes foram informados sobre os objetivos do estudo, os procedimentos envolvidos e o caráter voluntário da sua participação. Foi igualmente garantido que os participantes poderiam desistir da investigação em qualquer momento, sem qualquer tipo de consequência.

Antes da participação nas sessões dos *workshops*, foi obtido consentimento informado por escrito de todos os participantes envolvidos. Foram ainda asseguradas condições de confidencialidade e anonimato no tratamento dos dados recolhidos.

O estudo foi desenvolvido no âmbito do projeto InControl – Desenvolvimento centrado no utilizador e avaliação do custo-efetividade de um programa digital de gestão da dor para pessoas mais velhas, respeitando as normas institucionais e as boas práticas de investigação científica aplicáveis à área da saúde e da comunicação.

7 Considerações finais

A presente investigação parte da necessidade de compreender como as pessoas mais velhas com dor crónica se relacionam com diferentes conteúdos audiovisuais de apoio à execução de exercício físico, num contexto em que a saúde digital assume um papel cada vez mais relevante. Ao analisar a perceção desta população face a figuras humanas reais, avatares digitais e figuras criadas por inteligência artificial, este estudo pretende aprofundar o conhecimento sobre a influência da representação visual, das instruções e dos elementos audiovisuais na compreensão do exercício, na confiança, na motivação e na adesão à prática de exercício físico.

Espera-se que a investigação contribua para superar uma lacuna identificada na literatura, ao centrar-se especificamente em pessoas mais velhas com 65 ou mais anos com dor crónica e ao articular a tecnologia, a comunicação e a saúde. Este estudo tem uma forte relevância social, uma vez que poderá apoiar a criação de soluções

audiovisuais mais inclusivas, acessíveis e eficazes, promovendo a participação ativa das pessoas mais velhas em contextos digitais de saúde e de reabilitação.

Em última instância, pretende-se que os resultados reforcem o papel do audiovisual como ferramenta de comunicação em saúde, não apenas ao nível informativo, mas também enquanto mediador de confiança, empatia e envolvimento na prática de exercício físico.

Agradecimentos. Este trabalho foi realizado no âmbito do projeto InControl- Desenvolvimento centrado no utilizador e avaliação do custo-efetividade de um programa digital de gestão da dor para pessoas mais velhas, financiado pelo FEDER - Fundo Europeu de Desenvolvimento Regional através do COMPETE 2030 e por fundos Portugueses através da FCT- Fundação para a Ciência e a Tecnologia na framework do projeto COMPETE2030-FEDER-00781500 <https://doi.org/10.54499/2023.17304.ICDT>.

Referências

1. Moore, A.J., Richardson, J.C., Sim, J., Bernard, M., Jordan, K.P.: Older people's perceptions of remaining physically active and living with chronic pain. *Qualitative Health Research* **24**(6), 761–772 (2014). <https://doi.org/10.1177/1049732314529663>
2. Garofano, M., Del Sorbo, R., Calabrese, M., Giordano, M., Di Palo, M.P., Bartolomeo, M., Ragusa, C.M., Ungaro, G., Fimiani, G., Di Spirito, F., Amato, M., Ciccarelli, M., Pascarelli, C., Scanniello, G., Bramanti, P., Bramanti, A.: Remote rehabilitation and virtual reality interventions using motion sensors for chronic low back pain: A systematic review of biomechanical, pain, quality of life, and adherence outcomes. *Technologies* **13**(5), 186 (2025). <https://doi.org/10.3390/technologies13050186>
3. Wicks, M., Dennett, A.M., Peiris, C.L.: Physiotherapist-led, exercise-based telerehabilitation for older adults improves patient and health service outcomes: A systematic review and meta-analysis. *Age and Ageing* **52**(11) (2023). <https://doi.org/10.1093/ageing/afad207>
4. Čaić, M., Avelino, J., Mahr, D., Odekerken-Schröder, G., Bernardino, A.: Robotic versus human coaches for active aging: An automated social presence perspective. *International Journal of Social Robotics* **12**(4), 867–882 (2020). <https://doi.org/10.1007/s12369-018-0507-2>
5. Dino, M.J.S., Dion, K.W., Abadir, P.M., Budhathoki, C., Huang, C.M., Padula, W.V., Ong, I., Dennison Himmelfarb, C.R., Davidson, P.M., Thiamwong, L.: What drives older adults' acceptance of virtual humans? A conjoint and latent class analysis on virtual exercise coach attributes for a community-based exercise program. *Computers in Human Behavior* **164**, 108507 (2025). <https://doi.org/10.1016/j.chb.2024.108507>
6. Garcia, M.B.: Teachers in the metaverse: The influence of avatar appearance and behavioral realism on perceptions of instructor credibility and teaching effectiveness. *Interactive Learning Environments* **33**(7), 4334–4350 (2025). <https://doi.org/10.1080/10494820.2025.2462144>
7. Puri, A., Baker, S., Hoang, T.N., Zuffi, R.C.: To be (Me) or not to be? Photorealistic avatars and older adults. In: *ACM International Conference Proceeding Series*, pp. 503–507 (2017). <https://doi.org/10.1145/3152771.3156166>
8. Chen, T.L., Bhattacharjee, T., Beer, J.M., Ting, L.H., Hackney, M.E., Rogers, W.A., Kemp, C.C.: Older adults' acceptance of a robot for partner dance-based exercise. *PLoS ONE* **12**(10) (2017). <https://doi.org/10.1371/journal.pone.0182736>

9. Kamali, M.El, Angelini, L., Caon, M., Carrino, F., Rocke, C., Guye, S., Rizzo, G., Mastropietro, A., Sykora, M., Elayan, S., Kniestedt, I., Ziyhan, C., Lettieri, E., Khaled, O.A., Mugellini, E.: Virtual coaches for older adults' wellbeing: A systematic review. *IEEE Access* **8**, 101884–101902 (2020). <https://doi.org/10.1109/ACCESS.2020.2996404>
10. Lin, J.H.T., Wu, D.Y.: Exercising with embodied young avatars: How young vs. older avatars in virtual reality affect perceived exertion and physical activity among male and female elderly individuals. *Frontiers in Psychology* **12** (2021). <https://doi.org/10.3389/fpsyg.2021.693545>
11. Liu, B., Markopoulos, P., Tetteroo, D.: How anthropomorphism affects user acceptance of a robot trainer in physical rehabilitation. In: *HEALTHINF 2019 - Proceedings of the 12th International Conference on Health Informatics*, pp. 30–40 (2019). <https://doi.org/10.5220/0007343600300040>
12. Straßmann, C., Krämer, N.C., Buschmeier, H., Kopp, S.: Age-related differences in the evaluation of a virtual health agent's appearance and embodiment in a health-related interaction: Experimental lab study. *Journal of Medical Internet Research* **22**(4) (2020). <https://doi.org/10.2196/13726>

Documenting the Production Procedures of an In-Motion 360-Degree Video for Senior Virtual Cycling Tourism Experiences

Fabício Marques¹[0009-0007-9195-7818], Cláudia Pedro Ortet²[0000-0003-4113-2571], Liliana Vale Costa³[0000-0003-2451-3073], and Ana Isabel Veloso⁴[0000-0002-5070-0756]

DigiMedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal
{fabricio.marques, claudiaortet, lilianavale, aiv}@ua.pt

Abstract. The recent advancements in 360-degree video technology have made it a viable and cost-effective option for creating immersive virtual reality audiovisual content, increasing its application in areas such as healthcare, education, and entertainment. However, the specificities and procedures of 360-degree video production require further documentation, particularly concerning the use of action cameras and in-motion recordings. This paper addresses this gap by reporting findings of a Development Research (DR) study focused on the production of a 360-degree video designed for a senior virtual cycling tourism experience. It summarizes three expert interviews on the subject and details the procedures employed during the video production. The content was evaluated in a virtual cycling tourism experience with 22 participants aged 60 years and over, who provided feedback on motion aspects, image quality, and video perspective. The results highlight strategies to mitigate motion sickness, challenges in audio recording and post-production, and recommended image resolution for a satisfactory head-mounted display (HMD) viewing experience. These insights contribute to best practices in the production of immersive virtual reality content based on a 360-degree video.

Keywords: virtual reality; 360-degree video; cycling tourism.

1 Introduction

The rapid development of immersive technologies in recent years has established virtual reality (VR) as an attractive topic of research in several areas such as healthcare, education, and entertainment. In this context, 360-degree video emerges as a viable option for creating immersive content for VR experiences. This format offers a more cost-effective production process in comparison to VR created using computer-generated imagery (CGI), due to factors including reduced expenses and technical requirements [1].

The 360-degree video – also referred to as spherical video or panoramic video – is a medium that allows viewers to observe the image from all angles and directions. It can be watched on flat-panel displays (computers, tablets, or smartphones) or experienced

through a head-mounted display (HMD). When viewed with an HMD, a 360-degree video is commonly referred to as 360-degree VR – also known as 360° VR or 360 VR. This differentiation seems necessary due to the frequent association of VR with CGI environments. A defining characteristic that distinguishes 360-degree video from CGI-based VR is its use of real-world imagery [2]. This characteristic significantly influences the perceived realism of the content, offering a distinct experience compared to computer-generated simulations [3].

Studies suggest that 360-degree VR has the potential to enhance the level of immersion and the sense of presence [4, 5]. These factors can contribute to a higher perception of enjoyment and increased engagement with the content [6, 7], making it a valuable instrument for multiple purposes, such as serving as an assistive technology for seniors [8] or improving high school student learning outcomes [9].

Although 360-degree VR has its advantages, the video production presents unique challenges. For instance, motion sickness is the most referenced negative outcome from this format [10], demanding a series of measures to minimize its effect.

Despite the fact that several studies report on the feasibility, usability, and effects of 360-degree VR applications [5–7], the specificities and procedures of its production require further documentation. This is particularly relevant concerning the use of 360-degree action cameras and in-motion recordings, which increase challenges and make motion sickness concerns even more significant.

This paper addresses this gap by documenting the 360-degree video production procedures within a study focused on a senior virtual cycling tourism experience. It summarizes three expert interviews, details the strategies employed to manage the challenges of in-motion recording, and reports on participant feedback.

2 Methodology

The 360-degree video production documented in this paper was a central phase of a study based on a Development Research (DR) framework, aiming to develop immersive audiovisual content designed for a senior virtual cycling tourism experience. The DR approach generally has two emphases: developing practical and innovative ways to solve real problems and proposing general design principles to guide future decisions [11]. Based on the DR framework proposed by De Villiers (2005), this study was conducted in three phases: (i) problem analysis and video design, grounded in a literature review and a series of expert interviews; (ii) video production, focused on the development of a 360-degree video; and (iii) application and evaluation, in which the content was tested with the target population in an experience where participants rode a stationary bicycle while watching it on a flat-panel display (e.g., television) and using an HMD. Although this paper focuses on Phase 2, it also includes information from Phases 1 and 3 that contributed to the video production.

3 Production Procedures

3.1 Expert interviews

Three professionals and academics specialized in 360-degree video participated in semi-structured interviews focused on assessing their views on the benefits and challenges of the technology, obtaining technical recommendations to produce this content, and exploring their perspectives on the application in a virtual cycling tourism experience with seniors. The sample included a university professor in multimedia and communication technologies; a doctoral student and lecturer in audiovisual production with six years of experience with 360-degree videos; and a professional with three years of work in the production of 360-degree videos.

Experts highlighted immersion as the primary advantage of this technology. In parallel, the main challenge is to mitigate physical discomfort caused by motion sickness. As noted by one expert, this issue dominates 360-degree videos, particularly when experienced through an HMD, making it a priority to counteract.

Interviewees identified three main aspects influencing immersion and motion sickness: movement, audio, and lighting. Excessive image movement was cited as a critical factor. While current cameras feature stabilization, experts recommended additional measures to minimize unintended motion, such as using external stabilizers and avoiding rough routes or abrupt turns. Regarding audio, external recorder were suggested for better quality and post-production editing to enhance immersion. Lighting was also considered essential, as overexposed images can cause discomfort in an HMD. Furthermore, high image resolution was strongly emphasized, with experts suggesting at least 8K resolution for an optimal experience.

To conclude, interviewees provided insights on designing for senior citizens. To prevent falls, the viewer should be seated. Additionally, a slow editing rhythm was recommended, emphasizing to always start in a stationary position before initiating movement. These recommendations were crucial for the 360-degree video design and pre-production definitions described in the following section.

3.2 360-degree video production

The production of the 360-degree video was conducted between June and August 2024. As the objective was to create immersive content for a senior virtual cycling tourism experience, the video was composed of first-person cycling footage from tourist cycling routes in the city of Aveiro, Portugal. The work followed a three-stage video production method: pre-production, production, and post-production.

Pre-Production. This stage focused on logistics to anticipate challenges. A primary design choice was to segment the video into five distinct locations. While a continuous video was possible, this segmentation allowed for environmental variety. Based on literature and experts, location selection prioritized well-paved routes and straight paths to minimize footage movement and motion sickness. Also, the cycling pace was kept slow to moderate. The total video length was limited to between five and ten minutes.

Regarding equipment, the GoPro Max camera was selected for its availability and 5.6K resolution, satisfying the 4K minimum standard for HMDs. Despite the 8K recommendation, no camera with this capability was available for this research. To ensure a practical and reproducible setup that captures a cyclist's perspective, the camera was attached to the filmmaker's cycling helmet using a standard mount (see Fig. 1). For audio, a Zoom H4n external recorder was used to ensure professional quality. The recorder was strapped to the filmmaker's back to minimize wind noise.

The filming schedule was planned for the European summer to maximize the probability of favorable weather. Sessions were conducted in the morning, a time selected for optimal lighting conditions.

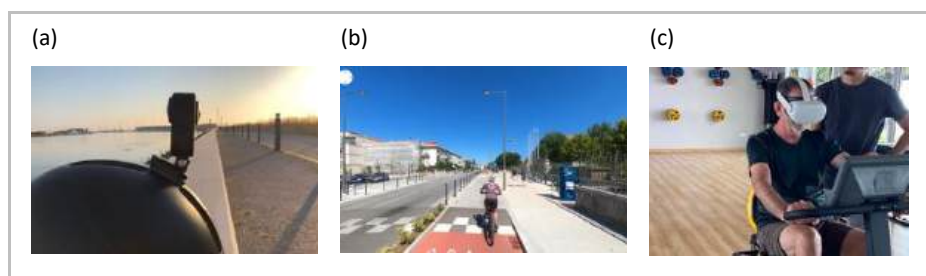


Fig. 1. Overview of production and evaluation phases: (a) the helmet-mounted camera setup; (b) a frame from the final 360-degree video; (c) a participant during the experience.

Production. This stage was conducted in three sessions. The first two involved only the filmmaker, while the third included a second cyclist. This aligned with the experience goal of evaluating the target population's preference for virtual cycling alone versus cycling with a partner.

As mentioned, the GoPro Max was used for filming with recording settings configured to 5.6K resolution and 30fps. Specific parameters included ISO 100, High Sharpness, and GoPro Color, while Shutter and White Balance were set to Auto. Audio was recorded using the Zoom H4n in stereo mode, with the WAV format at 48kHz and 16-bit.

Before filming, “room tone” was recorded at each location. This technique captured ambient sound without movement, ensuring cleaner audio. Approximately one minute of static footage was also recorded at the first location before cycling began.

At least three takes, each approximately five minutes long, were recorded per segment to ensure sufficient options for post-production. Particular attention was paid to maintaining a slow to moderate cycling pace to mitigate motion sickness. In the session involving the second cyclist, adjusting speed and maintaining spacing were challenging, necessitating two additional takes to establish the desired rhythm (see Fig. 1).

Post-Production. The final stage involved footage evaluation and editing. First, the 360-degree video was converted from the camera format to an editing format (.MOV) using GoPro Player software. Export settings were configured to 5.6K resolution,

equirectangular projection, CineForm codec, and Film Scan 1 quality, with both World Lock and Horizon Level activated. This configuration balanced image quality and file size, resulting in files averaging 15GB for two minutes of footage per location.

Converted files were imported into Adobe Premiere 2024, chosen for its native 360-degree editing tools. An appropriate sequence was created for VR footage, and the “Toggle VR Video Display” was activated. A “First Person” starting point was defined using the “Offset” effect, centering the initial view on the cyclist’s perspective.

The sequence begins with the static footage, followed by an introduction and the start of the cycling activity. For transitions between locations, a “Dip to Black” effect was applied, simulating a “long blink” to facilitate a smooth change of scene.

Audio was reconstructed in post-production as wind noise compromised the recordings made during movement. Only the “room tone” recorded in static positions had sufficient quality to be used, while specific sounds, such as boat motors or car engines, were integrated using authentic audio captured on location. Background music was added to three segments to evaluate user preferences, while the remaining two retained only ambient sound. Similarly, the second cyclist appears in three of the five segments to assess social preferences.

Narration was added to introduce the experience, describe location details, and conclude the video. Lastly, graphic information, including the project title and street names, was applied using a “Lower Third” template.

The video was exported using Adobe Media Encoder 2024 in HEVC (H.265) format. Export settings were configured to resolution of 5376×2688 and a frame rate of 29.976 fps, with Render at Maximum Depth and Maximum Render Quality activated. A VBR 1-pass bitrate targeting 100 Mbps was utilized. VR properties was set to Monoscopic, and audio was encoded in AAC format. The video was uploaded to YouTube as unlisted file for research purposes (https://youtu.be/Es_O5A0QV04).

3.3 Application and Evaluation

A convenience sample of 22 seniors (aged 60 and over) participated in the exploratory study, conducted between 12 and 31 August 2024. The participants were invited to ride a stationary bicycle while viewing the 360-degree video on a television and using an HMD (see Fig. 1).

The setup for the experience included a semi-recumbent stationary bicycle with a seat position adjustable according to the participant’s height. For the segment where participants watched the video on a flat-panel display, a 32-inch television was positioned one meter from the bicycle, connected through an HDMI cable to a laptop. For the VR HMD segment of the experience, the Oculus Meta Quest 2 was used. Although the 360-degree video was available on YouTube and could be viewed using the YouTube app on the Meta Quest 2, the internet connection available was not sufficient for an adequate video experience. As an alternative, the 360-degree video was uploaded to the Meta Quest 2 internal memory, a process made possible via the Windows system. This allowed the video to be viewed in the Meta Quest TV app in the highest possible resolution.

Post-experience feedback was collected using a questionnaire adapted from the brief version of the Senior Technology Acceptance Model (STAM) [12]. Additionally, participants took part in semi-structured interviews to assess overall perceptions, specific preferences, and recommendations for improvement. Data were analyzed using a mixed-methods approach: quantitative STAM scores were analyzed using descriptive statistics, while qualitative interview data were analyzed through a general inductive approach [13].

The brief version of the STAM instrument consists of a 14-item scale reflecting four factors: attitudinal beliefs, control beliefs, gerontechnology anxiety, and health. The questionnaire was administered using a 10-point Likert scale, where “1” represents “strongly disagree” and “10” represents “strongly agree”. The attitudinal belief factor showed a positive inclination among participants towards the technology presented, with a mean (M) of 8.0 and a standard deviation (SD) of 2.2. Regarding control beliefs (M=7.4; SD=3.0), results indicated that most participants felt capable of interacting with virtual cycling tourism, however, the technology was not perceived as highly accessible. Participants also reported minimal concerns in relation to the gerontechnology anxiety factor (M=2.6; SD=2.4). Regarding the health factor, participants generally indicated positive perceptions of their physical health and overall satisfaction with social and personal lives (M=8.7; SD=1.1).

Analysis of the semi-structured interviews revealed a generally positive response, with participants expressing high enthusiasm for the VR component. The immersion, realism, and sense of presence provided by the 360-degree VR were highlighted as primary advantages. For instance, the majority of participants (n=20) preferred the HMD over the television.

Regarding motion aspects, the design measures adopted to minimize motion sickness appeared successful. Most participants did not report any discomfort (n=17). The remaining five experienced only slight symptoms that did not significantly affect their experience. This suggests that the slow to moderate pacing, combined with the selection of well-paved, straight cycle paths, were successful strategies.

However, specific technical limitations were identified, primarily concerning image quality and video perspective. While participants enjoyed the content, some requested higher image resolution for better clarity. The camera on the filmmaker’s helmet also created a perspective that some users described as too high for a bicycle ride.

Despite these issues, the inclusion of narration and graphic information was valued by all participants. Overall, the experience was successful, with 15 seniors expressing willingness to adopt virtual cycling tourism as a regular or occasional activity, while the remaining seven considered the possibility at an older age.

4 Conclusion

This paper documents the complete production procedures of a 360-degree video developed specifically for a senior virtual cycling tourism experience. By detailing the pre-production, production, and post-production stages, it provides practical information for creating feasible in-motion immersive audiovisual content.

This research highlights specific strategies to mitigate motion sickness, which remains a primary concern in VR experiences. Data suggests that selecting locations with well-paved roads and straight paths to ensure footage stability, combined with a slow to moderate cycling pace, effectively minimizes physical discomfort. Additionally, beginning with a period in a static position before the movement allows users to adapt to the VR environment, enhancing comfort.

Regarding technical challenges, audio recording in motion proved difficult, making post-production reconstruction a reliable approach. Furthermore, while the 5.6K resolution satisfied the minimum requirements for a satisfactory HMD viewing, prioritizing cameras with 8K resolution or higher is recommended to maximize immersion. For cycling videos, it is fundamental to ensure that the camera captures an adequate first-person perspective.

These insights contribute to best practices in the production of immersive virtual reality content based on 360-degree video. It addresses, particularly, the unique challenges of in-motion filming designed for an HMD viewing experience.

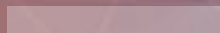
Acknowledgments. This work was supported by the IC Senior X project (GIP3_2022), FCT – Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia, I.P.), and DigiMedia – Digital Media and Interaction Research Center, under the project UIDB/05460/2020.

References

1. Jong, M.S.Y., Jiang, M.Y.C., Wong, P.S.P.: Pilot Study of Using ARCS Model to Investigate Motivational Affordances of EduVenture VR from Perspective of Construction Management Students. In: 5th IEEE Eurasian Conference on Educational Innovation 2022, ECEI 2022 (2022)
2. Snelson, C., Hsu, Y.C.: Educational 360-Degree Videos in Virtual Reality: a Scoping Review of the Emerging Research. *TechTrends*. 64, (2020). <https://doi.org/10.1007/s11528-019-00474-3>
3. Li, W., Qian, L., Feng, Q., Luo, H.: Panoramic video in education: A systematic literature review from 2011 to 2021, (2023)
4. Fonseca, D., Kraus, M.: A comparison of head-mounted and hand-held displays for 360° videos with focus on attitude and behavior change. In: *AcademicMindtrek 2016 - Proceedings of the 20th International Academic Mindtrek Conference* (2016)
5. Sundar, S.S., Kang, J., Oprean, D.: Being There in the Midst of the Story: How Immersive Journalism Affects Our Perceptions and Cognitions. *Cyberpsychol. Behav. Soc. Netw.* 20, (2017). <https://doi.org/10.1089/cyber.2017.0271>
6. Daltoè, T., Ruth-Herbein, E., Brucker, B., Jaekel, A.K., Trautwein, U., Fauth, B., Gerjets, P., Göllner, R.: Immersive insights: Unveiling the impact of 360-degree videos on preservice teachers' classroom observation experiences and teaching-quality ratings. *Comput. Educ.* 213, (2024). <https://doi.org/10.1016/j.compedu.2023.104976>
7. Lee, S.H., Sergueeva, K., Catangui, M., Kandaurova, M.: Assessing Google Cardboard virtual reality as a content delivery system in business classrooms.

- Journal of Education for Business. 92, (2017). <https://doi.org/10.1080/08832323.2017.1308308>
8. Ortet, C.P., Veloso, A.I., Vale Costa, L.: Cycling through 360° Virtual Reality Tourism for Senior Citizens: Empirical Analysis of an Assistive Technology. *Sensors*. 22, (2022). <https://doi.org/10.3390/s22166169>
 9. Yamwongsri, P., Chen, H.L.: The Effects of Spherical Video-Based Virtual Reality in Inquiry-Based Learning on Student Learning Outcomes, Higher-Order Thinking, and Learning Experiences in Geography Education. *Journal of Educational Computing Research*. 63, (2025). <https://doi.org/10.1177/07356331251360440>
 10. Häkkinen, J., Ohta, F., Kawai, T.: Time course of sickness symptoms with HMD viewing of 360-degree videos. In: *IS and T International Symposium on Electronic Imaging Science and Technology* (2019)
 11. de Villiers, M.R.: Three Approaches as Pillars for Interpretive Information Systems Research: Development Research, Action Research and Grounded Theory. In: *Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries*. pp. 142–151. South African Institute for Computer Scientists and Information Technologists, ZAF (2005)
 12. Chen, K., Lou, V.W.Q.: Measuring senior technology acceptance: Development of a Brief, 14-Item Scale. *Innov. Aging*. 4, (2020). <https://doi.org/10.1093/geroni/igaa016>
 13. Thomas, D.R.: A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*. 27, (2006). <https://doi.org/10.1177/1098214005283748>

Advanced Services and TV 3.0



Rede Pública: Portal de Integração para Conteúdos Audiovisuais e Serviços Digitais na TV 3.0 Brasileira

Valdecir Becker^[0000-0002-0639-3875], Wesley Alves^[0009-0000-1862-8670], Maristela Riquelme^[0009-0000-4497-6476], Clarissa Pereira^[0009-0008-0028-8979], Barbara Silva^[0009-0005-7178-4169] and Gabriel Aguiar^[0009-0006-1159-5291]

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

Abstract. A implementação da TV 3.0 no Brasil marca uma evolução na experiência audiovisual, oferecendo melhor qualidade e maiores possibilidades de interação. Este artigo apresenta o desenvolvimento da plataforma Rede Pública, que integra canais públicos brasileiros e serviços digitais governamentais em um portal instalado de fábrica em todos os receptores de TV comercializados no país. Utilizando Design Thinking, Design Science Research e prototipação em Figma, foram criados modelos para TV e aplicativo de segunda tela, visando engajamento, inclusão e acessibilidade. O projeto destaca o papel estratégico da TV 3.0 na cidadania digital e propõe, como próximos passos, ampliar testes com usuários para validar e fortalecer a integração entre Estado e sociedade.

Keywords: TV 3.0, E-Gov, Plataforma unificada, Interatividade.

1 Introdução

A televisão brasileira desempenha um papel significativo na comunicação nacional, evidenciando avanços e integrando o país desde sua inauguração em 1950, seguido pela adoção do sistema PAL-M, na década de 1970. Em 2007 começou a transição para o modelo digital, baseado na alta definição. A partir de 2010, foi incorporado o middleware Ginga, que acrescentou recursos de interatividade [1]. A próxima revolução deve ser a migração para uma TV 3.0, denominada comercialmente de DTV+, prevista para iniciar em 2026.

Neste modelo, a convergência entre internet (broadband) e transmissão tradicional (broadcast) torna-se robusta, viabilizando uma mescla total de funcionalidades. Diferenças como transmissão do sinal e interatividade entre as mídias desaparecem, tornando a origem do conteúdo transparente para o indivíduo. Esses fatores consolidam uma experiência ainda mais dinâmica e interativa para o público brasileiro.

O processo de transição para a TV 3.0 foi oficializado pelo Decreto N° 12.595, de 27 de agosto de 2025 [2], que instituiu também a Plataforma Comum de Comunicação Pública e Governo Digital. Conforme disposto no Artigo 19, essa plataforma integrará conteúdos e aplicações provenientes de entidades públicas, tendo como objetivos principais, conforme elencado no Artigo 20, ampliar a interação entre a população e as

instituições públicas, além de fomentar a inclusão digital e o acesso a serviços e informações governamentais. Esta plataforma deverá estar presente em todos os aparelhos de TV comercializados no país. O acesso aos fluxos de vídeo das emissoras poderá ser tanto por broadcast (quando o sinal estiver disponível no ar), ou por broadband (na ausência deste sinal). Desta forma, prevê-se uma cobertura nacional dos canais públicos federais. O acesso aos serviços de governo eletrônico e a atualização dos dados destes serviços será via broadband.

A comunicação pública se destaca como uma das áreas mais impactadas pela transição de modelos televisivos. Diversos órgãos do poder público – executivo, legislativo e judiciário – possuem canais próprios de televisão com diferentes níveis de abrangência. Da mesma forma, universidades e organizações da sociedade civil disponibilizam conteúdos audiovisuais por meio desse meio. A convergência e integração entre internet e televisão também potencializam os serviços de governo eletrônico (e-gov).

Diante desse contexto, este artigo apresenta a proposta da plataforma Rede Pública, um portal televisivo que centraliza a comunicação pública, abrangendo o Executivo Federal (incluindo canais da Empresa Brasileira de Comunicação e outras emissoras ministeriais), Poder Judiciário, Câmara dos Deputados, Senado Federal, emissoras legislativas estaduais e municipais, TVs universitárias e TVs comunitárias. O acesso ao portal é viabilizado por meio das TVs 3.0 e de um aplicativo de segunda tela, que oferece funcionalidades interativas e opções de navegação.

Sob essa perspectiva, a iniciativa objetiva desenvolver um ecossistema integrado, unificando conteúdos midiáticos, mecanismos de interação e serviços de governo digital. Os conceitos de design e as soluções de interação aplicados à plataforma digital proposta para o governo federal, juntamente ao aplicativo de segunda tela, foram desenvolvidos com base nas necessidades identificadas pelos diversos stakeholders envolvidos na comunicação pública nacional.

2 Fundamentos

A transição para a TV 3.0 insere a televisão aberta no paradigma dos aplicativos, já comum em smart TVs, marcando, dessa forma, a próxima fase da evolução audiovisual. Esse modelo híbrido não só melhora a qualidade visual e sonora, como também redefine a experiência do usuário, convertendo a televisão em uma plataforma interativa para serviços, conteúdos e novas relações entre produção e fruição audiovisual. No Brasil, essa transformação foi recentemente formalizada pelo Decreto nº 12.595, de 27 de agosto de 2025 [2], que instituiu as diretrizes para a implementação da TV 3.0 como etapa estratégica do Sistema Brasileiro de Televisão Digital (SBTVD). Esse processo se alinha também ao debate internacional sobre convergência midiática e cultura participativa [3-5]. É nesse cenário que se insere a proposta de Rede Pública, plataforma projetada para funcionar como um portal central que reúne, em um único local, canais de TV públicos, estatais, e uma ampla gama de serviços de e-gov.

A evolução do modelo de televisão digital baseado na alta definição e na interatividade mediada pelo Ginga para este novo paradigma de oferta de conteúdos e serviços

baseado em aplicativos representa mais do que um simples redimensionamento mercadológico de plataformas já existentes [1, 6]. Diferentemente de dispositivos como celulares e notebooks, a TV interativa, baseada em aplicativos, exige a adaptação a desafios específicos de usabilidade, atendendo tanto ao público que utiliza a televisão como principal meio de informação, quanto à parcela da sociedade que recorre a meios alternativos para se manter informada. Essa discussão converge com a noção de “cultura da convergência” de Jenkins [3] e com análises recentes da ITU sobre os novos ecossistemas digitais [7]. Aspectos como a manipulação do dispositivo por meio do controle remoto, a distância do usuário em relação à tela e a ausência de um cursor [8] assumem relevância nesse contexto. Embora o conteúdo broadcast possa atenuar parte dessas limitações, o projeto da TV 3.0 propõe uma abordagem mais abrangente, ao integrar dispositivos complementares como extensões do conteúdo, e não apenas como plataformas alternativas de acesso.

Com o passar dos anos, as inovações tecnológicas se tornaram aliadas em diversas tarefas executadas do dia a dia. O consumo de conteúdo audiovisual deixou de ser feito exclusivamente por uma tela (TV), passando a ocorrer em múltiplas plataformas, como celulares e computadores, em um movimento de adaptação ao surgimento de novas mídias. Diante dessa nova forma de consumir, consolidou-se o conceito de segunda tela (second screen) ou tela companheira (companion screen), um dispositivo adicional que permite ao usuário interagir com o conteúdo consumido. No contexto da TV 3.0, este conceito avança para uma integração sincronizada e bidirecional, onde o smartphone deixa de ser um dispositivo paralelo para tornar-se uma extensão interativa da experiência broadcast. Estudos de D’heer e Courtois [9], Holmes et al. [10] e Celestino e Pernisa Júnior [11] reforçam a importância da segunda tela para o engajamento, destacando que a sincronia entre o fluxo principal e os conteúdos extras é condição fundamental para a imersão do telespectador.

3 Métodos

O desenvolvimento das aplicações e da plataforma começou com workshops de levantamento de requisitos, realizados em parceria com órgãos do governo federal responsáveis pela comunicação, informação e prestação de serviços ao público, tanto presencialmente quanto por meio digital. Esses encontros seguiram a metodologia do Design Thinking [12,13], que valoriza empatia, trabalho multidisciplinar e prototipagem iterativa, facilitando a identificação e organização dos requisitos de acordo com as variadas demandas sociais e perfis de cidadania dos envolvidos. Nessas reuniões também foi escolhido o nome provisório “Rede Pública”, que o governo federal pode alterar futuramente.

Na sequência, foram elaborados protótipos no Figma [14] do portal instalado nas TVs e do aplicativo de segunda tela, visando incorporar os requisitos definidos. Nesta etapa adotou-se o método da Design Science Research (DSR) [15,16], que guia o desenvolvimento de artefatos a partir da definição do foco e do problema a ser solucionado por software. A DSR também guia a validação destes artefatos, tanto internamente, quanto com os stakeholders envolvidos na fase de elicitação de requisitos. Todo

o processo de design e prototipação foi realizado de forma colaborativa e validado junto aos órgãos públicos, garantindo atendimento às necessidades institucionais e resultando em um Produto Mínimo Viável (MVP) com clareza visual, navegação intuitiva e interatividade alinhada aos objetivos do projeto. As versões finais foram aprovadas pelos órgãos envolvidos na definição inicial dos requisitos.

4 Design e desenvolvimento

A metodologia de design e desenvolvimento adotada para a concepção do Portal Rede Pública e suas aplicações foi baseada em um processo de prototipagem de interface. O principal objetivo foi a análise e o teste dos requisitos do projeto, bem como a organização da arquitetura do sistema. Conforme Decreto, o aplicativo da Rede Pública deve vir instalado nativamente nos receptores, bastando ao usuário apenas a atualização. Dessa forma, pretende-se possibilitar o acesso aos canais públicos e estatais a todos os receptores, mesmo em lugares onde não há sinal terrestre desses canais.



Fig. 1. Tela inicial.



Fig. 2. Estrutura em carrosséis.

Um conceito central do portal Rede Pública é o acesso e a navegação por camadas, tornando a experiência mais complexa e personalizada à medida que o indivíduo interage com o sistema. Dessa forma, a TV linear é utilizada como ponto de partida para este novo universo digital, atendendo pessoas com pouca, ou nenhuma, alfabetização digital. Ao selecionar o aplicativo da Rede Pública é reproduzido um vídeo com destaques das emissoras, com a opção de acessar diretamente a emissora ou pular para o Portal. Em ambas as opções, a tela inicial do Portal (Fig. 1) apresenta uma lista horizontal, navegável pelo controle remoto, de canais lineares. Há uma transição suave entre o vídeo de abertura e o primeiro canal, que ocupará a tela inteira da TV automaticamente caso não haja interação. A navegação entre os canais poderá ser feita tanto pela interface aplicativo, quanto pelas teclas de navegação do controle remoto. Dessa forma, aproxima-se a experiência baseada em aplicativos com a cultura do *zapping*.

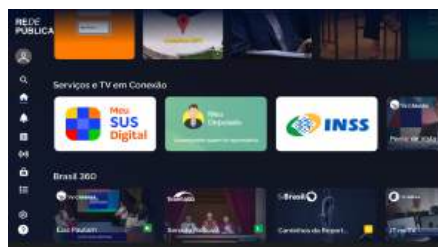


Fig. 3. Navegação por serviços de e-gov.

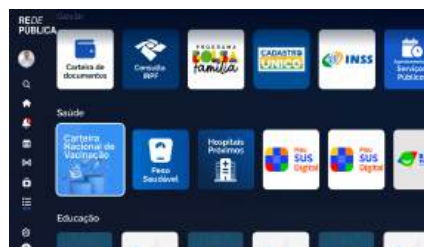


Fig. 4. Meu Espaço.

O portal é baseado em carrosséis horizontais, com acesso aos canais ao vivo no topo (Fig. 2). No lado esquerdo há um menu, com acessos alternativos, de configuração e de personalização. Conforme a navegação se aprofunda, descendo nos menus do carrossel, são apresentados conteúdos sob demanda dos canais (definidos pela curadoria desses canais), recomendações complementares aos canais (definidas com base na experiência do usuário). Posteriormente, há uma transição para a oferta de serviços de e-gov, mesclando inicialmente conteúdos de TV com aplicativos, para carrosséis compostos apenas por serviços de e-gov (Fig. 3). Finalmente, há uma seção, denominada Meu Espaço, totalmente personalizada a partir de um login feito no aplicativo de segunda tela (Fig. 4).

4.1 Aplicativo de segunda tela

Além do Portal, a Rede Pública também conta com um aplicativo mobile, a fim de proporcionar mais interatividade e mobilidade à experiência do usuário (Fig. 5, 6 e 7). Na tela inicial, os conteúdos são complementares à programação televisiva, seguindo a estrutura da plataforma.

Na parte inferior, há o menu com cinco atalhos definidos por padrão (Fig. 5): Início, que dá acesso à página inicial; Guia de Programação, que apresenta a programação linear das emissoras públicas; Controle remoto, que emula as funções do controle físico, completo, da TV e permite a navegação tanto pelo portal Rede Pública quanto pelo aplicativo DTV+, possibilitando também a interação enquanto o usuário alterna entre os canais lineares; Modo Secundário, que oferece acesso a documentos disponibilizados pelo gerador de conteúdo, enquetes e formulários, por exemplo; e Acessibilidade, que agrega os recursos de configuração da acessibilidade, abrangendo tanto os obrigatórios por norma do SBTVD, quanto os específicos da aplicação e do aplicativo e do Portal. Esses atalhos, entretanto, podem ser personalizados pelos usuários de acordo com as preferências de uso.



Fig. 5. Tela inicial.

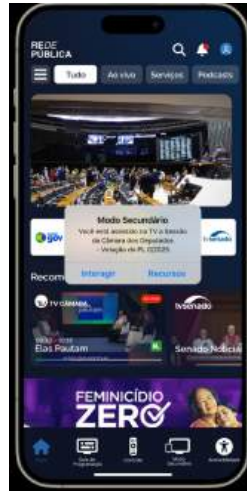


Fig. 6. Modo secundário.



Fig. 7. Busca.

Ainda na tela principal, destaca-se a busca integrada, que permite que, ao assistir a um conteúdo na TV e acionar a lupa no aplicativo, o usuário receba automaticamente sugestões relacionadas ao tema exibido, sem precisar digitar nada (Fig.7). Caso queira expandir a busca, é possível digitar palavras-chave, escolher temas definidos por tags generalistas ou clicando em um determinado conteúdo, o que possibilita pesquisas cruzadas em todo o Portal, incluindo os canais ao vivo, conteúdos sob demanda e e-gov.

Este aplicativo estará disponível nos sistemas de download dos fabricantes de celulares, com a chancela do governo federal. Como trata-se de um aplicativo de segunda tela, ele é complementar ao Portal, acrescentando informações, interações e navegações que podem ser complexas utilizando o controle remoto na TV. Para o acesso a estes recursos, o Portal Rede Pública deverá estar aberto na TV. Sem isso, não há a sincronização dos recursos entre portal e segunda tela.

5 Conclusões e trabalhos futuros

A proposta da plataforma Rede Pública evidencia como a TV 3.0 pode se consolidar como um recurso estratégico para ampliar a interação entre Estado e sociedade, integrando os conteúdos das emissoras públicas e serviços de e-gov. A solução apresentada responde a desafios como a baixa usabilidade dos sistemas interativos baseados apenas no controle remoto [17] e se apoia na ideia de adaptação contínua da televisão frente às novas mídias [8].

Ao propor uma arquitetura multiplataforma, o trabalho demonstra que a TV interativa não deve ser entendida apenas como evolução tecnológica, mas como oportunidade de fortalecimento da cidadania digital, pavimentando também um caminho para que outras redes televisivas nacionais tenham um aparato para se embasar, fortalecendo a

soberania tecnológica brasileira. O Portal Rede Pública, articulado ao aplicativo de segunda tela, oferece uma experiência mais inclusiva, pautada em princípios de acessibilidade e usabilidade, e alinhada à lógica de consumo multiplataforma.

Para trabalhos futuros, pretende-se ampliar de forma significativa o escopo dos testes com usuários reais do sistema, contemplando uma amostra diversificada da população brasileira em termos de faixa etária, nível de escolaridade, condições socioeconômicas e familiaridade com tecnologias digitais. Essa ampliação tem como finalidade avaliar de maneira mais abrangente a experiência de uso da aplicação, garantindo que sua concepção esteja alinhada aos princípios de usabilidade, acessibilidade e inclusão. Dessa forma, busca-se assegurar que a plataforma seja eficaz não apenas para os grupos que já utilizam recursos digitais cotidianamente, mas também para aqueles que encontram maiores barreiras no acesso a tecnologias, contribuindo, assim, para a redução das desigualdades no acesso à informação e aos serviços digitais.

Referências

1. Becker, V. (2013). A evolução da interatividade na televisão: da TV analógica à era dos aplicativos. *Lumina*, 7(2).
2. BRASIL. (2025) “Decreto nº 12.595, de 27 de agosto de 2025”. Diário Oficial da União, Brasília. <https://www.in.gov.br/en/web/dou/-/decreto-n-12.595-de-27-de-agosto-de-2025-651425519>.
3. Jenkins, H. (2006) “Convergence Culture: Where Old and New Media Collide”. New York: NYU Press.
4. Chernock, R., Gomez-Barquero, D., Whitaker, J., Park, S. I., & Wu, Y. (2016). ATSC 3.0 next generation digital TV standard—An overview and preview of the issue. *IEEE Transactions on Broadcasting*, 62(1), 154-158.
5. Rabaça, R. S., de Oliveira, G. H. M. G., Junior, L. N. D. S., Lopes, P. B., Park, S. I., & Akamine, C. (2024, June). Evaluation of ATSC 3.0 MODCODs for TV 3.0. In 2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB) (pp. 1-5). IEEE.
6. Lim, B. M., Park, S. I., Ahn, S., Seo, J. H., & Hur, N. (2024). Features and Applications of Physical Layer Multiplexing and MIMO in ATSC 3.0. *SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING*, 10.
7. ITU. (2018) “Guidelines for the Transition to Digital Television Broadcasting”. ITU Publications.
8. Scolari, C. A. (2008) “Hacia la hipertelevisión: Los primeros síntomas de una nueva configuración del dispositivo televisivo”. *Revista Académica de la Federación Latinoamericana de Facultades de Comunicación Social. Diálogos de La Comunicación*, Nº 77, Julio, diciembre.
9. D’heer, E. and Courtois, C. (2016) “The changing dynamics of television consumption in the multimedia living room”. *Convergence*, v. 22, n. 1, p. 3–17.
10. Holmes, D. et al. (2017) “Second screen and TV audience engagement: A systematic literature review”. *Journal of Broadcasting & Electronic Media*, v. 61, n. 2, p. 220–239.
11. Celestino, R. R. and Pernisa Júnior, C. A. (2021) “Televisão interativa e convergência digital: perspectivas para a TV 3.0 no Brasil”. Rio de Janeiro: Appris.
12. Plattner, H., Meinel, C., & Leifer, L. (2010). “Design Thinking: Understand–Improve–Apply”. Springer.

13. Brown, T. (2009) “Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation”. Harper Business.
14. FIGMA, Inc. Figma. São Francisco, CA: Figma, 2025. <https://www.figma.com>
15. Dresch, A., Lacerda, D. P., & Junior, J. A. V. A. (2020). Design science research: método de pesquisa para avanço da ciência e tecnologia. Bookman Editora.
16. Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European journal of information systems*, 25(1), 77-89.
17. Canatta, F. (2014) “TV e Segunda Tela: uma análise do horário nobre no Twitter”.
18. Dissertação (Mestrado em Comunicação Social) – Faculdade de Comunicação Social – PUCRS, Porto Alegre. <http://repositorio.pucrs.br/dspace/handle/10923/5648>

TV Senado na Era da TV 3.0: Uma Proposta de Experiência do Usuário para o Acesso à Informação Legislativa

Maristela de Freitas Riquelme , Clarissa Evelin Pereira ,
Wesley Alves da Silva  e Valdecir Becker 

Laboratório de Interação e Mídia, Centro de Informática, Universidade Federal da
Paraíba, João Pessoa, Paraíba, Brazil
contato@lim.ci.ufpb.br

Abstract. Este artigo apresenta o desenvolvimento de uma proposta de inserção da TV Senado no modelo da TV 3.0. Essa nova geração integra conteúdos de radiodifusão e banda larga, ampliando a oferta audiovisual e os serviços interativos para os cidadãos. Para lidar com essa convergência, o projeto seguiu diretrizes de Experiência do Usuário (UX), priorizando acessibilidade, usabilidade e engajamento. A metodologia aplicou a abordagem de Design Thinking em oficinas de elicitação de requisitos com as equipes técnicas e de gestão da emissora. Protótipos interativos foram desenvolvidos no Figma e validados ao longo do processo, apoiando decisões relacionadas à hierarquia da informação, usabilidade e design gráfico. Como resultado, foi produzido um Produto Mínimo Viável (MVP) funcional que integra a transmissão legislativa ao vivo a recursos interativos de navegação e contextualização das atividades parlamentares, ampliando o acesso às informações do Senado Federal. A solução também fornece uma referência para futuras iniciativas de televisão interativa no Brasil. Trabalhos futuros incluem testes com perfis diversos de usuários, refletindo a variabilidade social, econômica e educacional do país, a fim de refinar a experiência e fortalecer a conexão entre o público e o Poder Legislativo.

Keywords: TV 3.0 · Televisão Legislativa · Experiência do Usuário · Televisão Interativa

1 Introdução

A TV Senado, pioneira na transmissão legislativa de alcance nacional no Brasil desde 1996, tem como missão constitucional dar transparência aos atos parlamentares e aproximar o cidadão do debate político [1]. Atualmente, o principal desafio da emissora reside na natureza unidirecional de sua transmissão [2]. Embora as sessões plenárias e comissões sejam transmitidas ao vivo, o espectador permanece passivo, sem ferramentas integradas para aprofundar o conhecimento sobre as pautas ou interagir com o rito legislativo em tempo real. A transição

para a TV 3.0 apresenta-se como a solução para esse cenário, pois permite converter o sinal de vídeo em um canal bidirecional, reunindo conteúdos e aplicações de entidades públicas para ampliar o acesso a informações do Legislativo [3].

Diante dessa transformação tecnológica, este artigo investiga a evolução da experiência do usuário no consumo da informação legislativa. O objetivo é superar a limitação técnica do *broadcast* tradicional ao integrar funcionalidades que possibilitem ao cidadão acompanhar a tramitação de projetos e a atuação de seus representantes sem sair da interface da televisão. A TV 3.0 viabiliza essa transformação ao permitir que aplicativos rodem sobrepostos ao sinal de vídeo, utilizando a banda larga para fornecer dados personalizados e serviços de governo digital diretamente ao controle remoto.

Neste contexto, este trabalho apresenta uma proposta de interface interativa para a TV Senado alinhada ao paradigma da TV 3.0. A solução integra a transmissão linear das sessões legislativas a camadas interativas adicionais de informação e navegação, permitindo ao usuário acessar conteúdos contextuais, acompanhar parlamentares específicos e explorar dados legislativos relacionados às transmissões. A proposta foi materializada em um protótipo funcional que demonstra como a convergência entre radiodifusão e banda larga pode ampliar a transparência e a compreensão da atividade parlamentar.

2 Diretrizes do design centrado no usuário

O Decreto nº 12.595/2020 [3] institui o paradigma da TV 3.0 baseada em aplicativos, ampliando a qualidade audiovisual e incorporando serviços interativos. Nesse cenário, o alinhamento da TV Senado ao novo modelo exige clareza, previsibilidade e acessibilidade, e o *design* centrado no usuário oferece o arcabouço para orientar decisões ao priorizar necessidades reais dos usuários em vez de uma lógica exclusivamente tecnológica [4].

A experiência do usuário (UX) envolve percepções e respostas ao uso de um sistema, e a usabilidade — pilar da IHC — representa a facilidade com que os usuários atingem seus objetivos [5]. Em ambientes audiovisuais digitais e plataformas de mídia interativa, o design de interface exerce papel fundamental na forma como os usuários exploram conteúdos e serviços, influenciando diretamente engajamento, navegação e compreensão das funcionalidades disponíveis [6]. A ISO 9241-11 [7] define usabilidade por eficácia, eficiência e satisfação, enquanto Nielsen [8] acrescenta aprendizado, memorização, erros e segurança, fatores essenciais em contextos televisivos marcados por controle remoto, distância de visualização e ausência de cursor.

O design centrado no usuário, conforme Norman [9] e a ISO 9241-210 [10], orienta o desenvolvimento pela compreensão de tarefas e contextos, com participação contínua e avaliação iterativa, sustentando a definição de fluxos e padrões visuais na TV 3.0. Diretrizes como as heurísticas de Nielsen [8] e as regras de Shneiderman [11] garantem interfaces claras e previsíveis, enquanto o *User Experience Honeycomb* de Morville [12] amplia essa análise ao incluir dimensões como desejabilidade, credibilidade, encontrabilidade e valor.

A arquitetura da informação [13][14][15] organiza e rotula conteúdos para assegurar navegação eficiente em ecossistemas amplos. A acessibilidade, guiada pelas diretrizes do W3C [16], fundamentou decisões sobre contraste, tipografia e navegação simplificada, garantindo inclusão desde a concepção.

Assim, as diretrizes de *Design Centrado no Usuário* (DCU) [17], abordagem que posiciona as necessidades dos usuários como fundamento das decisões projetuais, permeiam todo o projeto, permitindo desenvolvimento orientado por evidências, reduzindo atritos e assegurando à TV 3.0 uma plataforma eficiente, acessível e alinhada às demandas do ecossistema audiovisual em transformação.

3 Metodologia

A proposta para inserção da TV Senado no ecossistema da TV 3.0 foi estruturada para integrar rigor técnico, alinhamento institucional e uma abordagem centrada no usuário. O processo iniciou-se com um diagnóstico técnico para identificar como os conteúdos da emissora poderiam ser integrados ao novo modelo baseado em aplicativos.

A elicitação de requisitos ocorreu em oficinas colaborativas fundamentadas no *Design Thinking*, que contaram com a participação de membros do corpo diretivo da emissora, apresentadores e secretários técnicos. Diferente de plataformas voltadas ao entretenimento, as prioridades identificadas nas oficinas focaram na fidedignidade da informação e na rastreabilidade da atividade parlamentar. Essa diretriz orientou a definição de funcionalidades voltadas ao acompanhamento da atividade parlamentar, explorando a integração entre o sinal de *broadcast* e dados provenientes do portal institucional do Senado. Essa composição permitiu consolidar requisitos que orientaram a prototipagem iterativa, desenvolvida com validação contínua junto à equipe da TV Senado.

A concepção da interface priorizou a redução da carga cognitiva, adaptando princípios de IHC à navegação por controle remoto. Por meio de protótipos no Figma, simulou-se a experiência baseada em aplicativos, incluindo transmissões legislativas ao vivo, conteúdos institucionais sob demanda, serviços informativos e recursos de acessibilidade. As decisões estéticas respeitaram a identidade visual do Senado Federal, utilizando uma tipografia adequada à leitura em TV e um layout modular inspirado no Portal Unificado das emissoras legislativas. O resultado foi um MVP funcional que demonstra a viabilidade da proposta para a transparência ativa no ambiente televisivo.

4 Resultados e discussão

A proposta apresentada foi concebida especificamente para o contexto editorial e institucional da TV Senado, buscando ampliar sua função histórica de transparência legislativa por meio das possibilidades interativas introduzidas pelo paradigma da TV 3.0.

A partir das etapas de ideação e validação conduzidas com a equipe da emissora, foram definidos requisitos de experiência que orientaram o desenvolvimento

da interface. Entre eles destacam-se: a continuidade audiovisual da transmissão ao vivo durante a navegação, de modo que o conteúdo legislativo permaneça sempre acessível ao usuário; a redução da carga cognitiva na exploração das informações, obtida por meio de uma organização clara dos conteúdos; a eficiência de navegação entre transmissões ao vivo, conteúdos sob demanda e dados contextuais; a rastreabilidade das atividades parlamentares, permitindo que o usuário acompanhe atores políticos; e a fidedignidade das informações apresentadas, assegurando que os dados exibidos estejam alinhados às fontes oficiais do Senado Federal.

A análise das contribuições obtidas nas oficinas resultou em uma proposta de interface voltada à transformação da experiência de consumo da TV Senado. A solução propõe a transição de um modelo essencialmente unidirecional para um ecossistema de transparência ativa. Enquanto na televisão linear convencional o espectador ocupa um papel predominantemente passivo diante da programação *broadcast*, a evolução para a TV 3.0 permite que a emissora passe a operar também como uma interface bidirecional. Nesse cenário, a arquitetura da aplicação utiliza um modelo baseado em aplicativos para sobrepor, em tempo real, dados institucionais ao sinal de vídeo, atendendo principalmente aos requisitos de fidedignidade da informação e rastreabilidade parlamentar.

Com base nesses princípios, a proposta de navegação foi estruturada a partir de um menu lateral principal, responsável por organizar os serviços e categorias de conteúdo. Essa decisão de design busca facilitar a identificação rápida das funcionalidades e reduzir o esforço cognitivo do usuário, mantendo os elementos essenciais visíveis durante a exploração da interface.

A Figura 1 apresenta duas telas da aplicação, nas quais é possível observar como os requisitos definidos foram materializados na experiência do usuário.

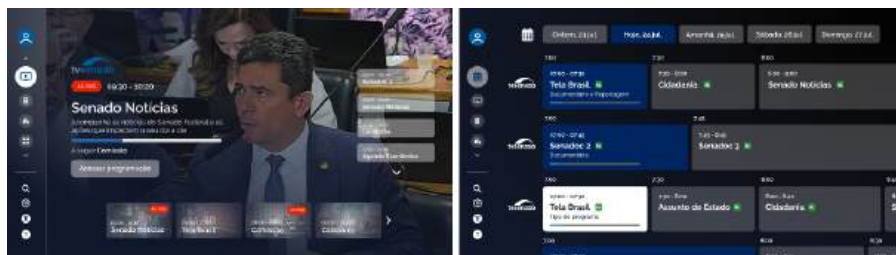


Fig. 1. Telas de TV ao vivo e EPG [Autorial]

Na tela à esquerda, observa-se o menu lateral associado à experiência de TV ao vivo. Nesse ambiente, são reunidos o Guia Eletrônico de Programação (EPG), transmissões ao vivo, conteúdos em vídeo, materiais promocionais em formato vertical, vídeos sob demanda (VOD) e serviços institucionais oferecidos pelo Senado Federal. Para garantir fluidez na interação dentro do modelo *lean-back*, o

fluxo principal inicia-se na TV ao vivo. A transmissão permanece em reprodução ao fundo enquanto um carrossel de canais é exibido na parte inferior da tela, permitindo alternar rapidamente entre transmissões sem interromper o consumo. Essa solução atende diretamente ao requisito de continuidade audiovisual.

À direita, o foco é o Guia Eletrônico de Programação (EPG), cujo acesso foi simplificado para atender aos requisitos de eficiência e redução da carga cognitiva. O acionamento pode ser realizado ao pressionar a seta para cima no ícone de TV ou diretamente a partir da transmissão ao vivo, oferecendo múltiplos caminhos para a mesma funcionalidade. A interface do guia apresenta uma consulta detalhada com filtros por data e horário, proporcionando uma visão abrangente que facilita a memorização e o aprendizado do sistema.

Na Figura 2 há a continuação das funcionalidades propostas.

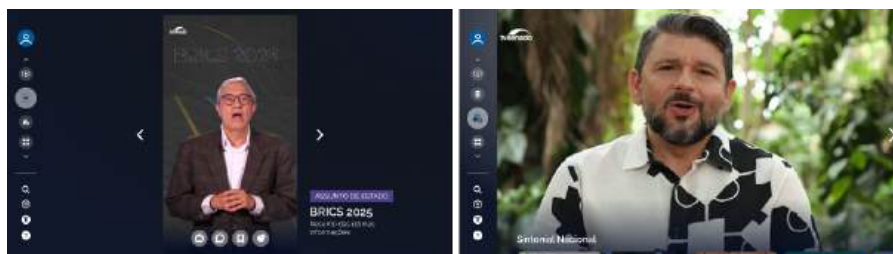


Fig. 2. Telas de Conteúdo vertical e VOD [Autorial]

À esquerda, encontra-se o terceiro item do menu, denominado Conteúdo Vertical. Inspirado em formatos contemporâneos de consumo audiovisual em dispositivos móveis, esse recurso organiza conteúdos legislativos em sequências curtas e navegáveis. O engajamento é incentivado por meio de ações contextuais, como curtir, comentar e salvar, acessíveis diretamente pelos comandos de seta sobre o próprio conteúdo.

Por fim, à direita, apresenta-se a área dedicada aos vídeos promocionais. Nessa seção, os conteúdos são organizados em uma navegação horizontal ampla, priorizando destaques institucionais e facilitando a descoberta de novos materiais audiovisuais.

A Figura 3 apresenta outra funcionalidade relevante da aplicação.

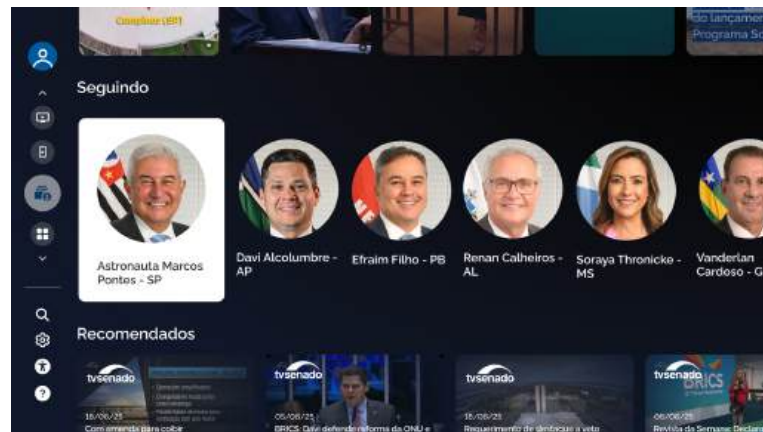


Fig. 3. Tela de Seguindo [Autorial]

Nela, é exibida a funcionalidade “Seguindo”, que permite ao usuário acompanhar políticos e outras figuras públicas do legislativo. Esse recurso torna a personalização explícita, ao possibilitar a consulta a agendas, projetos e demais ações associadas aos parlamentares, fortalecendo a transparência e o vínculo institucional com o cidadão. Diferentemente da televisão linear convencional, essa camada de interação utiliza o canal de retorno *broadband*, permitindo que o espectador atue como curador de seu próprio fluxo informacional.

5 Conclusão e trabalhos futuros

Os resultados demonstram que a proposta alinhada ao modelo de TV 3.0 para a TV Senado integra princípios de interatividade, acessibilidade e usabilidade, revelando oportunidades em um cenário de transformação da mídia. Ao viabilizar a curadoria ativa da informação e o acesso direto a serviços, o modelo supera a passividade do espectador tradicional e evidencia o potencial da televisão como ferramenta de participação democrática e inclusão digital. Nesse contexto, o MVP desenvolvido demonstra a viabilidade de integrar a transmissão legislativa ao vivo a recursos interativos, ampliando o acesso às informações do Senado Federal e contribuindo para transformar a experiência da TV legislativa.

Como este ciclo inicial de desenvolvimento concentrou-se na validação técnica e institucional junto à equipe da emissora, trabalhos futuros devem priorizar a realização de testes de usabilidade com diferentes perfis de usuários em escala nacional. A inclusão de amostras populacionais diversas permitirá avaliar a efetividade da interface em contextos reais de uso e orientar novas iterações de design. Dessa forma, será possível aprimorar a experiência oferecida e garantir que a solução atenda às necessidades de públicos com diferentes perfis sociais, educacionais e regionais.

Em síntese, o projeto apresenta um caminho possível para a evolução da TV Senado no contexto da TV 3.0, demonstrando como a convergência entre radiodifusão e serviços digitais pode ampliar a transparência legislativa e fortalecer o acesso da sociedade às atividades do Poder Legislativo.

Referências

1. TV Senado: TV Senado completa 30 anos: veja um pouco dessa história (2026). <http://www.youtube.com/watch?v=nALj8C2586U>, último acesso em 11/03/2026
2. Fórum SBTVD: TV 3.0 Project, Standard. Fórum SBTVD, São Paulo, BR (2020). https://forumsbtvd.org.br/tv3_0/#, último acesso em 3/12/2025
3. BRASIL: Decreto nº 12.595, de 27 de agosto de 2025. Diário Oficial da União, Brasília (2025). <https://www.in.gov.br/en/web/dou/-/decreto-n-12.595-de-27-de-agosto-de-2025-651425519>, último acesso em 20/11/2025
4. Garrett, J. J.: The elements of user experience: user-centered design for the web and beyond. 2 ed. New Riders, Berkeley (2011)
5. Preece, J., Rogers, Y., Sharp, H.: Interaction design: beyond human-computer interaction. 5 ed. Wiley, Hoboken (2019)
6. Soto-Sanfiel, M. T., Aymerich-Franch, L.: Designed to binge? Exploring user perceptions of interface features on video streaming platforms. Acta Psychologica. Elsevier, Amsterdam (2025)
7. ISO 9241-11:2018: ergonomics of human-system interaction – Part 11: usability: definitions and concepts. ISO, Geneva (2018)
8. Nielsen, J.: Usability engineering. Academic Press, Boston (1993)
9. Norman, D. A.: The design of everyday things. Rev. and expanded edn. Basic Books, New York (2013)
10. ISO 9241-210:2019: ergonomics of human-system interaction – Part 210: human-centred design for interactive systems. ISO, Geneva (2019)
11. Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., Diakopoulos, N.: Designing the user interface: strategies for effective human-computer interaction. 6th edn. Pearson, Boston (2017)
12. Morville, P.: User experience honeycomb (2004). https://semanticstudios.com/user_experience_honeycomb, último acesso em 26/11/2025
13. Rosenfeld, L., Morville, P., Arango, J.: Information architecture: for the web and beyond. 4 ed. O'Reilly, Sebastopol (2015)
14. Wurman, R. S.: Information anxiety. Doubleday, New York (1989)
15. Resmini, A., Rosati, L.: Pervasive information architecture: designing cross-channel user experiences. Morgan Kaufmann, Burlington (2011)
16. World Wide Web Consortium: Web Content Accessibility Guidelines (WCAG) 2.1. Recommendation, 05 June 2018. <https://www.w3.org/TR/WCAG21/>, último acesso em 26/11/2025
17. Lowdermilk, T.: User-Centered Design: A Developer's Guide to Building User-Friendly Applications. O'Reilly Media, Sebastopol (2013)

Automatic System for Reception, Processing, and Subtitling of ISDB-Tb One-Seg Signals for Hearing Accessibility

Gonzalo Olmedo¹, Victor Zurita¹, Bryan Castillo¹ and Nelson Benavides¹

¹ Universidad de las Fuerzas Armadas ESPE, Wicom-Energy, Departamento de Eléctrica, Electrónica y Telecomunicaciones, Sangolquí, Ecuador

[gfolmedo, nbbenavides1]@espe.edu.ec

Abstract. This paper presents the development and implementation of an automatic system for the reception, processing, and subtitling of Digital Terrestrial Television (DTT) signals, aimed at improving access to audiovisual content for people with hearing impairments. The system is based on the ISDB-Tb standard, widely adopted in Latin America, and employs open technologies and artificial intelligence tools. It focuses on the One-Seg service (Layer A), characterized by low bandwidth requirements and robustness in real-world environments. The proposed solution builds upon a previously developed ISDB-T receiver, which was enhanced for this research. Signal reception was carried out using GNU Radio and an ADALM-PLUTO software-defined radio (SDR) device. The reconstructed MPEG-TS transport stream was segmented using FFmpeg for interval-based processing. Each segment was processed to extract the audio, which was automatically transcribed to text through a speech recognition API. The resulting transcriptions were converted into subtitle files with precise timecodes and embedded into the videos through controlled re-encoding. Finally, the subtitled segments were merged into a single continuous playback file. The system operates autonomously and asynchronously, achieving an average latency of approximately five seconds plus the segment duration, making it suitable for near real-time applications. This proposal contributes to digital inclusion by facilitating equitable access to televised information for hearing-impaired audiences.

Keywords: Accessibility, Inclusion, DTT, ISDB-T, One-Seg, MPEG-TS, FFmpeg, API, Automatic Subtitling, SDR, GNU Radio.

1 Introduction

Digital Terrestrial Television (DTT) enables the transition from analog to fully digital broadcasting, offering improved audiovisual quality, spectral efficiency, and robustness through MPEG-2/MPEG-4 compression and OFDM modulation with forward error correction [1]. Ecuador adopted the ISDB-Tb standard in 2010 due to its strong performance in urban propagation, mobility scenarios, and support for enhanced digital services [2], [3].

ISDB-T employs a hierarchical structure that divides a 6 MHz channel into 13 segments, allowing HDTV, SDTV, and One-Seg mobile services to coexist within the same multiplex. The Brazilian variant (ISDB-Tb) incorporates H.264/AVC and HE-AAC codecs, improving compression efficiency, while the MPEG Transport Stream (TS) organizes audio, video, and metadata through 188-byte packets and PSI/SI tables such as PAT, PMT, SDT, and NIT [4], [5].

One-Seg, designed for mobile reception, operates on a single segment using robust parameters—QPSK modulation, FEC 2/3, and deep time interleaving—providing resilience under mobility and multipath fading [6]. Its widespread adoption makes it appropriate for applications requiring low bandwidth, high robustness, and real-time subtitling.

Software-Defined Radio (SDR) platforms such as the ADALM-PLUTO, which features wide frequency coverage and an FPGA-enabled SoC, allow implementing fully programmable ISDB-Tb receivers [7], [8]. This flexibility supports the development of processing pipelines needed to extract and manipulate One-Seg transport streams.

In Ecuador, accessibility regulations mandate subtitling and other services; however, compliance remains limited, creating barriers for deaf and hard-of-hearing audiences [9]. Automatic Speech Recognition (ASR) systems offer a practical alternative for generating real-time captions. Modern transformer-based ASR platforms, such as AssemblyAI, provide accurate transcriptions with timestamped outputs suitable for synchronized subtitles [10]. Recent studies show that ASR-based captioning can achieve error rates appropriate for accessibility applications [11], supporting the feasibility of the system proposed in this work.

2 Methodology

The proposed system comprises a complete processing pipeline that integrates software-defined radio (SDR) reception, real-time transport stream (TS) segmentation, audio extraction, speech transcription, and subtitle embedding. The workflow is designed to operate continuously, enabling near-real-time generation of captions for ISDB-Tb One-Seg broadcasts.

2.1 ISDB-Tb Receiver Implementation

Signal acquisition was performed using an ADALM-PLUTO SDR configured for ISDB-Tb One-Seg reception. GNU Radio was used to construct the demodulation chain, including synchronization, OFDM demodulation, symbol extraction, and transport stream reconstruction.

A custom GNU Radio processing block, **CC_Processor**, was included to isolate the One-Seg packets (segment 7), assemble them into a continuous transport stream, and forward them to the automated processing module. The receiver was validated under local broadcast conditions, demonstrating stable recovery of audio and video services.



Fig. 1. illustrates the reception system used in this work, showing the signal path from the SDR to the MPEG-TS generation.

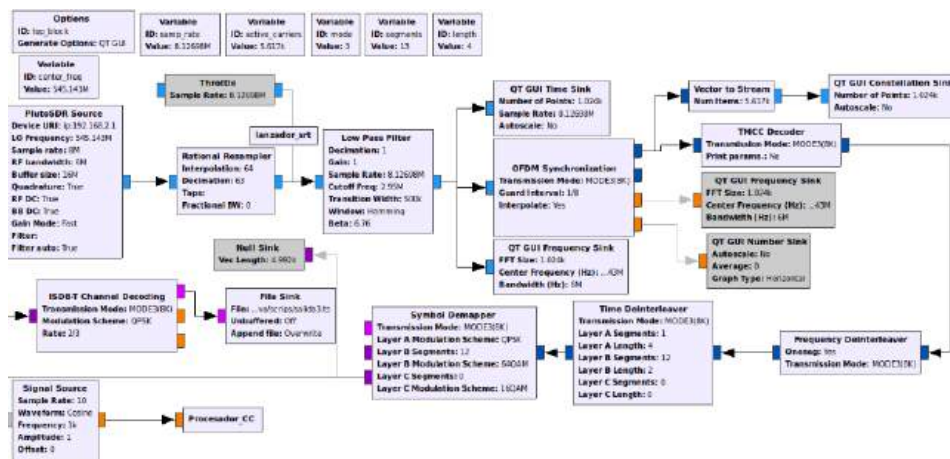


Fig. 2. Block diagram of the ISDB-T receiver in GNU Radio, highlighting the custom block "CC_Processor" for subtitle processing.

2.2 Automated Processing Block

The automated processing module, **Procesador_CC**, orchestrates the complete subtitling workflow. It receives the reconstructed transport stream and manages the following tasks TS segmentation, Audio extraction, ASR-based transcription, Subtitle file generation and Subtitle embedding and output storage.

By coordinating these operations, **Procesador_CC** ensures continuous processing in parallel with ongoing signal reception.

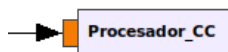


Fig. 3. shows the *Procesador_CC* block, which encapsulates

2.3 Real-time Transport Stream Segmentation

The output TS file (salida.ts) generated by the receiver grows continuously as the signal is received. To enable real-time processing and parallel workflows, the system divides the stream into fixed 5-second segments using FFmpeg.

The parameters used for the continuous TS reading are summarized in Table 1.

Table 1. Parameters used in the tail command.

Parameter	Description
-f	Reads the file as it grows.
-c +0	Reads from the beginning of the file.
salida.ts	Input MPEG-TS file generated by the SDR receiver.

This approach creates a series of small, time-aligned MPEG-TS files, enabling efficient audio extraction, transcription, and subtitle embedding with minimal latency.

2.4 Audio Extraction and ASR Integration

For each segment, audio was extracted using FFmpeg and converted into a standardized WAV file suitable for speech recognition. The extracted audio was then sent to an ASR system capable of generating text transcriptions and confidence scores.

Transcription outputs were processed to create synchronized subtitle files. Temporal alignment was achieved by associating each transcription with the corresponding 5-second video segment, ensuring accurate and consistent caption timing.

For speech recognition, the system relied on the AssemblyAI cloud-based ASR platform, which uses pre-trained transformer models optimized for high-fidelity audio-to-text conversion.

The transcription process follows a “word heard, word written” principle, meaning that the ASR output preserves the acoustic realization of the speech without replacing words based on contextual inference.

In addition, the API provides post-processing features such as automatic punctuation, formatting, and speaker diarization, which facilitate the generation of structured subtitle files.

2.5 Subtitle Generation and Embedding

Following transcription, subtitle files were created in .srt format and embedded into their corresponding TS segments using FFmpeg.

Processing times remained within the expected bounds, supporting real-time operation. The system handled uninterrupted segmentation and caption generation, demonstrating robust file management and processing stability.

3.3 ASR Performance and Subtitle Quality

The ASR system produced coherent transcriptions of news-broadcast audio, with confidence levels sufficient for subtitle generation. The pipeline successfully processed each segment's audio and converted it into synchronized text.

Subtitle quality reflected accurate alignment with spoken content, and the segmentation strategy proved compatible with real-time transcription and captioning scenarios.

To quantitatively evaluate transcription quality, the ASR outputs were compared against manually verified references using subtitle files in .srt format generated during the experiments.

A ground truth dataset of approximately 1600 words was created through manual transcription of the Spanish news-broadcast content. For this dataset, the AssemblyAI-based pipeline achieved a word accuracy of approximately 94.7%, corresponding to a Word Error Rate (WER) of about 5.3%.

The residual errors were mainly associated with background noise, unclear pronunciation, and overlapping speech, which is consistent with typical ASR limitations in real broadcast environments.

In order to contextualize the chosen ASR solution, several platforms were evaluated under the same dataset and segmentation conditions. Table 3 summarizes the average processing time per 5-second segment and the subscription model for each system.

Table 2. ASR platforms comparison – Processing performance.

Name	Processing	Avg. time	Subscription
AssemblyAI	Cloud	4.86 seconds	NO
Whisper	Local	10 seconds	NO
Deepgram	Cloud	13 seconds	YES
Vosk	Local	19 seconds	NO

Table 3 presents word accuracy and WER comparison results.

Table 3. ASR platforms comparison – Accuracy results (Spanish).

System	Word Accuracy	Word Error Rate (WER)
AssemblyAI	~94.7%	~5.3%
Whisper	~84.4%	~15.6%
Deepgram	~88.3%	~11.7%
Vosk	~90.1%	~9.9%

These results confirm that AssemblyAI offers the best trade-off between transcription accuracy and processing latency for the considered scenario, while also avoiding the need for high-end local hardware. In contrast, Whisper and Vosk require significantly more computational resources to run locally, which increases system cost and complexity despite their competitive accuracy.

3.4 Subtitle Embedding and Playback

The subtitles were successfully embedded into the One-Seg video segments, demonstrating correct integration of textual content into the audiovisual stream. After processing, the system concatenated the captioned segments into a single file (*buffer.ts*), which allowed uninterrupted playback.

The pipeline effectively managed file handling, directory organization, and buffered writing, ensuring smooth output generation.



Fig 6. Playback of content with subtitles

3.5 Latency Measurements

The system incorporated a time-measurement mechanism based on the `time()` function from Python's `time` module, renamed as `tiempo_actual`, enabling precise monitoring of execution times for key processes. This method was applied to tasks such as audio extraction, API upload, transcription, and subtitle embedding. Performance evaluation was conducted using 60 continuous segments derived from 40 audiovisual streams—20 approximately 3 minutes long and 20 ranging between 1 and 3 minutes. Table 4 presents the average processing times for 10-second and 5-second segments.

Conversion of the transcription into `.srt` format was not included in the measurements, as AssemblyAI performs this operation automatically and the script only adjusts the timing units, resulting in negligible duration. The presentation time depends on the length of the first segment, since the system operates in parallel: while the first segment is being displayed, the next one is already being received or processed. This pipelined behavior ensures continuous playback without increasing the overall system latency.

Table 4. Parameters used in the tail command.

Process Part	10-seconds segment	5-second segment
Audio extraction-MP3	0.27 seconds	0.22 seconds
Upload to AssemblyAI	1.34 seconds	1.25 seconds
Transcription	2.76 seconds	2.75 seconds
Subtitle embedding	0.84 seconds	0.64 seconds
Total	5.21 seconds	4.86 seconds

4 Conclusions

This work successfully demonstrated a complete SDR-based ISDB-Tb subtitling system capable of near real-time operation using open-source tools and cloud-based ASR. The automated pipeline reliably processed One-Seg signals, producing synchronized subtitles with 4.86 s average latency and 94.7% word accuracy (WER 5.3%) on a 1600-word Spanish broadcast corpus. These results confirm that low-cost SDR platforms combined with modern AI transcription services enable accessible digital broadcasting.

User validation studies with deaf participants, essential for assessing comprehension and usability, await ethics committee approval per Declaration of Helsinki standards. The current system provides technically validated automatic transcription and timing alignment as a proof-of-concept foundation for accessible subtitling.

Future work will incorporate linguistic adaptation for deaf audiences, speaker/sound event marking, ethically-approved user validation, multilingual support, lower-latency ASR, error correction, and emergency alert integration to fully realize the accessibility potential demonstrated by this research.

References

1. Lins, A.: Digital Television Systems Overview (2024)
2. Jácome, J.R., Pazmino, A.O., Gavilanes, J.J., Martínez, O.G., Ribadeneira, J.A.: Analysis of propagation models for digital terrestrial television (DTT), under the ISDB-Tb standard in cities of Ambato and Latacunga, Ecuador. *Journal of Science and Research*, 3(CITT2017), 84-90 (2018).
3. MINTEL: Plan Maestro de Transición a la Televisión Digital Terrestre (2018)
4. Pisciotto, J.: ISDB-T International Standard Overview (2010)
5. Avaro, O., Eleftheriadis, A., Herpel, C., Rajan, G., Ward, L.: MPEG-4 systems: overview. *Signal Processing: Image Communication*, 15(4–5), 281–298 (2000)
6. Paderna, C., Higashino, T., Okada, M.: ISDB-T One-Seg Performance Analysis (2014)
7. Mitola, J.: Software Radio Architecture (2002)
8. Ortiz, J., Zavala, D., Sosa, R.: Evaluation of the ADALM-PLUTO SDR Platform (2018)
9. García-Prieto, V., Figuereo-Benítez, J.C.: Accesibilidad de los contenidos televisivos para personas con discapacidad (2022)
10. AssemblyAI: API Documentation (2025)
11. Kuhn, K., Kersken, V., Zimmermann, G.: Communication Access Real-Time Translation Through Collaborative Correction of ASR. In: CHI EA '25 (2025)

Usability and UX Evaluations



Development of a Controlled Robotic Prototype for Fatigue Recognition Through Microexpression

Nancy Paredes¹, Alex Chiguano², Sofia Chicaiza² and Mayra Comina²

¹ Department of Electrical, Electronic and Telecommunications, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador, WiCOM-Energy Research Group

² Department of Energy and Mechanical, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

Corresponding author: niparedes@espe.edu.ec

Abstract. This paper presents the development and evaluation of a low-cost, modular robotic prototype designed to detect user fatigue through the recognition of facial microexpressions. The system combines a mechanical platform featuring pan-tilt head movement, synchronized arm gestures, and simulated blinking, with a computer vision model based on transfer learning using YOLOv5. The dataset, derived from Roboflow and labeled according to the Facial Action Coding System (FACS), was augmented and categorized into four key expressions: awake, tired, distracted, and yawning. The final model achieved a validation accuracy of 94.9%, recall of 93.6%, and mAP@50 of 96.2%. Experimental evaluation with 200 real-time predictions confirmed an overall accuracy of 78.25% and average precision of 56.5%, particularly excelling in distinguishing between "awake" and "tired" states. The integration of mechanical and vision-based feedback allows the robot to respond with expressive gestures, enhancing human-robot interaction. This work demonstrates the feasibility of using microexpression analysis for real-time fatigue monitoring in collaborative environments such as education, healthcare, or industrial supervision.

Keywords: Robot, microexpressions, recognition, fatigue.

1 Introduction

The advancement of collaborative robotics has redefined the parameters of interaction between humans and machines in industrial, healthcare, and educational settings. This evolution has been driven by the industry 5.0 paradigm, which promotes cooperation between human intelligence and artificial intelligence to create safer, more ergonomic, and personalized work environments [1]. In this context, collaborative robots (cobots) have become key tools due to their ability to operate in shared spaces without the need for physical barriers, enabled by the integration of sensors, adaptive control systems, and advanced perception technologies [2].

One of the most promising developments in this field is the incorporation of computer vision to interpret subtle human signals, such as facial microexpressions. Although brief and involuntary, these expressions convey relevant information about the user's emotional and physiological state, making them particularly useful for detecting fatigue, distraction, or drowsiness [3]. Automated recognition of microexpressions has gained special relevance in critical tasks such as operator supervision or driver monitoring, where fatigue-related errors can lead to severe consequences.

Detecting microexpressions has traditionally posed a technical challenge due to their low intensity and millisecond duration. However, the use of lightweight convolutional neural networks and models such as YOLOv5 has proven effective for real-time detection with high accuracy, even under variable lighting or angle conditions [4]. In parallel, the physical design of cobots has evolved toward modular, 3D-printed structures optimized for the integration of sensors, cameras, and low-cost actuators [5].

This paper presents the design and validation of a robotic prototype capable of detecting signs of fatigue through facial microexpressions. Both the mechanical construction of the system and the training of a YOLOv5-based computer vision model are described, with the goal of delivering a functional, accessible system suitable for human-robot collaboration scenarios.

2 State of art

The development of collaborative robots has transformed human-machine interaction, with safer work environments being among its top priorities. These systems rely on sensors, adaptive control, and computer vision to interpret subtle human signals and respond in real time [1], [6]. The detection of fatigue in operators is a key application, as this physiological state directly affects productivity and safety in industrial, educational, and transportation

contexts [7]. The analysis of facial microexpressions has established itself as a key resource to evaluate emotional and physiological states. Although its low intensity and duration, measured in milliseconds, make detection difficult, its coding using the Facial Action Coding System (FACS) allows gestures to be segmented into Action Units (AUs), facilitating their identification using computer vision algorithms [8]. Recent studies have shown that automatic detection of microexpressions can be a reliable predictor of fatigue, especially in monitoring drivers and operators [9-10].

The evolution of computer vision techniques has enabled the application of deep learning to enhance facial and microexpression recognition. Among these, transfer learning has become an effective strategy by leveraging models pre-trained on large databases, such as COCO, to specialize them on specific fatigue detection tasks [11-12]. In particular, the YOLO (You Only Look Once) family of algorithms has positioned itself as one of the most efficient methods in real-time detection. Optimized versions of YOLOv5 have demonstrated high accuracy in scenarios with limited computational resources and have been applied to the analysis of anomalous expressions and drowsiness [4,10].

Wu [13] developed an automated driver fatigue and distraction detection system using computer vision and machine learning, combining facial landmarks with lightweight convolutional networks to enable real-time inference. Similarly, Majdi [14] proposed DriveNet, a deep convolutional architecture integrated with a Random Forest classifier, achieving 95% accuracy in detecting distracting behaviors, such as computer use. Arava [6] integrated YOLOv5 with 3D facial landmarks to detect driver drowsiness, ensuring robustness. Kanigoro [15] combined facial landmarks and YOLOv5 to identify prolonged blinking in simulated driving environments, demonstrating the feasibility of these approaches for industrial applications. Onososen [16] implemented a YOLOv8-based system for fatigue detection in construction workers.

Ji [9] proposed a computer vision model to detect mental fatigue in chemical control room operators, integrating deep neural networks with random forests. Viquez [10] examined fatigue in industrial manufacturing environments using visual analysis, highlighting the usefulness of these tools for anticipating occupational hazards. Fang [17] investigated

the detection of muscle fatigue, supporting the integration of these systems into cobots capable of monitoring the worker's physical condition in real time. Taken together, the scientific literature confirms that fatigue detection using computer vision and microexpressions is a rapidly expanding field.

3 Methodology

This research was developed around the design, construction, and integration of a functional robotic system focused on detecting signs of fatigue through the analysis of facial microexpressions. The methodology was structured into three main phases: mechanical design of the prototype, the development of the computer vision system, and the integration and implementation of the complete system.

3.1 Mechanical Design of the Prototype

Its design was based on derived functional requirements, such as head movement with two degrees of freedom, basic arm mobility with one degree of freedom per arm, a simulated blinking mechanism for aesthetic purposes, and a modular, spacious structure to house electronic components, wiring, and the camera. Autodesk Inventor was used as the 3D modeling tool, allowing validation of interference, movement ranges, and ease of assembly [18], as shown in Fig 1.



Figure 1. CAD model of the robotic prototype

During the mechanical conceptualization process, morphological analysis and alternative generation techniques were employed to compare, comparing different mechanisms for each movable part [19].

The design of the robotic arms followed a modular and minimalist approach, assigning one degree of freedom to each arm. This configuration, commonly used in low-cost industrial systems, was sufficient to enable basic communicative gestures, such as raising the arms, which serve as a visual alert when drowsiness is detected in the user [20]. The parts were designed to be detachable, facilitating maintenance and allowing for potential future modifications.

The simulated blinking was implemented using a crank-slider mechanism which converts the rotation of a servomotor into vertical movement of the eyelids. Although its primary function is aesthetic, this element contributes to the robot's overall expressiveness.

3.2 Computer Vision Model

Two approaches were developed for detecting facial microexpressions related to fatigue. The first approach relied on a proprietary dataset captured under controlled laboratory conditions, where participants were exposed to tasks designed to induce visual fatigue. OpenFace [21] was employed to extract facial landmarks, head pose, gaze estimation, and AUs based on the FACS. These features were subsequently fed into machine learning classifiers, such as Support Vector Machines (SVM) and Random Forests, to discriminate between states of alertness and fatigue. Although this approach provided valuable insights, it faced notable limitations: the restricted dataset reduced the model's ability to generalize across

diverse populations, lighting conditions, and head orientations. Moreover, hyperparameter optimization required significant computational resources and manual tuning, which hindered scalability for real-time applications. Similar challenges have been reported in literature, where small datasets often limit the effectiveness of AU-based fatigue detection systems [22].

The second computer vision model, implemented in the robotic prototype, addressed these limitations by leveraging large-scale annotated datasets. For this purpose, images were retrieved from Roboflow Universe, a platform that builds upon the COCO dataset [23], which is widely used in object detection and recognition tasks. A subset of 928 images containing human faces was carefully selected, focusing on microexpressions associated with fatigue, distraction, and yawning [24]. These images were manually annotated on the Roboflow platform to highlight key facial regions such as eyes, eyebrows, and mouth, which are critical for identifying fatigue-related AUs. The annotation process followed established guidelines from FACS [25], ensuring that each facial region was coded in alignment with the muscular activations involved in drowsiness. To increase model robustness, preprocessing steps included geometric normalization, histogram equalization for illumination invariance, and resizing the images to 640×640 pixels. In addition, data augmentation techniques, the dataset was expanded to 2,217 annotated samples, including 1,936 training images and 281 validation images, enhancing variability. This dataset construction pipeline enabled the training of a YOLOv5-based model capable of detecting both macro-expressions

sions and subtle micro-expressions with real-time performance. The images were organized into four main categories: awake, tired, yawning, and distracted, based on identifiable microexpression patterns. The segmentation process focused on critical facial regions such as the eyes, eyebrows, and mouth, deliberately excluding the nose due to its low relevance in detecting fatigue-related expressions. Annotation of the dataset was carried out using the Roboflow platform [25].

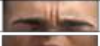
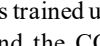
Each category was characterized according to AUs defined by the FACS. The “awake” category corresponded to a neutral facial expression, with no significant activation, primarily identified as AU0. The “tired” category was associated with combinations of AUs including AU41 to AU44, indicating drooping, narrowing, and eyelid closure; AU4, reflecting eyebrow lowering; AU15, for downturned lip corners; and AU17 and AU26, related to chin elevation and depression. The “yawning” category was primarily identified by AU27, corresponding to a wide-open mouth, in conjunction with AU43 and AU5, indicating eye closure and elevated eyelids, as illustrated in Table 1. Meanwhile, the “distracted” category was associated with AU5 and AU45, reflecting frequent blinking and raised eyelids [8].

3.3 Transfer Learning Training in Roboflow

The model was trained in the cloud using Roboflow Train, leveraging transfer learning with a YOLOv5 model pre-trained on the COCO dataset. This approach enables the reuse of previously learned features, such as face detection, thereby facilitating the learning of fatigue-related facial microexpressions [12].

The base layers of the model were frozen, and only the upper layers were retrained using the new dataset. This strategy reduced computational costs while improving the model's accuracy. The microexpression-based fatigue detection model was evaluated through an experiment involving 200 controlled predictions. During testing, researchers alternated facial expressions corresponding to the states “tired,” “awake,” “yawning,” and “distracted,” under well-lit conditions and at a maximum distance of 80 cm from the camera.

Table 1. Facial Action Coding System 3.0 [9]

Category	AU	Description	Example	Category	AU	Description	Example
Normal / Awake	AU0	Neutral face, with no significant activation of AUs		Yawn	AU26	Chin drop	
	AU41	Drooping of the upper eyelids			AU27	Mouth opening	
Tired	AU42	Narrowing of eyelid aperture			AU43	Eyes closed during the yawn	
	AU43	Eyes fully closed			AU5	Eyelid elevation	
	AU44	Marked narrowing of the eyelid aperture		Distracted	AU6	Upper eyelid elevation	
	AU4	Brow lowering			AU45	Flinching/blink	
	AU15	Lip corner depressor			AU24	Lip press	
	AU17	Chin raise					

4. Results

The artificial intelligence model was trained using transfer learning in Roboflow, leveraging the YOLOv5 pre-trained model and the COCO dataset. During internal validation, the model achieved strong performance, with a precision of 94.9%, a recall of 93.6%, and a mAP@50 of 96.2%. These metrics indicate a high ability to detect facial expressions associated with fatigue.

To further assess its performance, the model was evaluated using a multiclass confusion matrix, which enabled detailed analysis of accuracy and precision for each expression category. The classes that were tired and awake demonstrated the highest performance, reaching 78,25% accuracy and 56,5% precision. This suggests that the model is particularly effective in identifying the two categories most relevant to fatigue monitoring. Conversely, the distracted and yawning classes exhibited lower performance, with accuracies ranging from 75% to 78% and precisions between 50% and 56%, revealing challenges in distinguishing more subtle or overlapping expressions.

Overall, the results demonstrate the viability of the proposed system for human–robot interaction scenarios requiring non-intrusive fatigue monitoring. While further refinements are necessary, especially to improve discrimination of ambiguous expressions such as distraction the findings support the prototype’s potential for assistive, supervisory, and alerting applications.

5. Discussion

The results from the development of the robotic prototype demonstrate the feasibility of combining mechanical expressiveness with a computer vision system for fatigue detection via facial microexpressions. The synchronization between facial state detection and the robot’s mechanical response enhances its real-time interaction capabilities, which are essential in collaborative applications. The model’s performance decreased when classifying more subtle expressions, such as “distracted” and “yawning,” suggesting that, although effective, the dataset could benefit from more samples to improve the system’s robustness across varying facial, cultural, and environmental conditions. Furthermore, although the system proved functional in controlled tests, its evaluation in real environments remains a significant challenge. Factors such as variable lighting, spontaneous expressions, and non-simulated behavior may affect both the model’s accuracy and the system’s reliability. Despite these limitations, the proposed approach

successfully integrates mechanical and computational components within a coherent human–robot interaction framework, establishing a solid foundation for future implementations and technological improvements.

References

1. Rahman, Md. Mijanur, Khatun, Fatema, Jahan, Ismat, Devnath, Ramprosad, Bhuiyan, Md. Al-Amin, Cobotics: The Evolving Roles and Prospects of Next-Generation Collaborative Robots in Industry 5.0, *Journal of Robotics*, 2024, 2918089, 22 pages, 2024. <https://doi.org/10.1155/2024/2918089>.
2. Y Wang, Y., Wang, J., Feng, J. et al. Integrated task sequence planning and assignment for human–robot collaborative assembly station. *Flex Serv Manuf J* 35, 979–1006 (2023). <https://doi.org/10.1007/s10696-022-09479-2>.
3. J. Zhang, "Human-computer Interaction for Driver Fatigue Detection through Micro-expressions," 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE), Jinzhou, China, 2023, pp. 208-212, doi: 10.1109/ICSECE58870.2023.10263425.
4. Yao, K., Wang, Z., Guo, F., & Li, F. (2024). Driver Abnormal Expression Detection Method Based on Improved Lightweight YOLOv5. *Electronics*, 13(6), 1138. <https://doi.org/10.3390/electronics13061138>.
5. L. Čehovin Zajc, A. Rezelj and D. Skočaj, "Low-Cost Open-Source Robotic Platform for Education," in *IEEE Transactions on Learning Technologies*, vol. 16, no. 1, pp. 18-25, 1 Feb. 2023, doi: 10.1109/TLT.2022.3224227.
6. Arava, M., & Sundaram, D. M. (2024). Integrating lightweight YOLOv5s and facial 3D keypoints for enhanced fatigued-driving detection. *PeerJ. Computer science*, 10, e2447. <https://doi.org/10.7717/peerj-cs.2447>
7. Ji, Z., Xie, X., Jiang, E., Wang, Y., Min, B., Yang, S., Chen, Y., & Pons, D. (2025). Integrating DRN-RF with computer vision for detection of control room operator's mental fatigue. *PloS one*, 20(4), e0320780. <https://doi.org/10.1371/journal.pone.0320780>
8. A. Freitas-Magalhaes, *Facial Action Coding System 3.0 (F-M FACS 3.0)*, 1st ed., Elsevier, 2018. ISBN: 9789898766915
9. Ji Z, Xie X, Jiang E, Wang Y, Min B, et al. (2025) Integrating DRN-RF with computer vision for detection of control room operator's mental fatigue. *PLOS ONE* 20(4): e0320780. <https://doi.org/10.1371/journal.pone.0320780>
10. Alfavo-Viquez, D., Zamora-Hernandez, MA., Azorín-López, J. et al. Visual analysis of fatigue in Industry 4.0. *Int J Adv Manuf Technol* 133, 959–970 (2024). <https://doi.org/10.1007/s00170-023-12506-7>
11. Jirak, D., Belgiovine, G., Eldardeer, O., Sciutti, A., Rea, F. (2025). Vision-Based Muscular Fatigue Detection for Safe Human-Robot Collaboration: Evaluation of a Pilot Study. In: Palinko, O., et al. *Social Robotics. ICSR + AI 2024. Lecture Notes in Computer Science*, vol 15563. Springer, Singapore. https://doi.org/10.1007/978-981-96-3525-2_33
12. Ultralytics, "YOLOv5." GitHub. [En línea]. Disponible en: <https://github.com/ultralytics/yolov5> [Accedido: 06-dic-2025].
13. D. Wu, "Improving automatic detection of driver fatigue and distraction using machine learning," *arXiv*, arXiv:2401.10213, Jan. 2024. <https://doi.org/10.48550/arXiv.2401.10213>
14. M. S. Majdi, S. Ram, J. T. Gill and J. J. Rodríguez, "Drive-Net: Convolutional Network for Driver Distraction Detection," 2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Las Vegas, NV, USA, 2018, pp. 1-4, doi: 10.1109/SSIAI.2018.8470309.

15. B. Kanigoro, B. Asdyo, "Facial Landmark and YOLOv5 Drowsiness Detection System", *Procedia Computer Science*, Vol 245, pp 548-554, 2024, <https://doi.org/10.1016/j.procs.2024.10.281>.
16. Onososen, A. O., Musonda, I., Onatayo, D., Saka, A. B., Adekunle, S. A., & Onatayo, E. (2025). Drowsiness Detection of Construction Workers: Accident Prevention Leveraging Yolov8 Deep Learning and Computer Vision Techniques. *Buildings*, 15(3), 500. <https://doi.org/10.3390/buildings15030500>
17. C. Fang, L. Petemel, A. Seth, M. Sartori, K. Mombaur and E. Yoshida, "Human Modeling in Physical Human-Robot Interaction: A Brief Survey," in *IEEE Robotics and Automation Letters*, vol. 8, no. 9, pp. 5799-5806, Sept. 2023, doi: 10.1109/LRA.2023.3296349.
18. R. H. Shih, "Leaming Autodesk Inventor 2024: Modeling, Assembly and Analysis," SDC Publications, 2024. ISBN: 978-1-63057-586-1
19. M. Nathan et al., "Design and Testing of a Pan-Tilt Mechanism for Severe Environments," MIT DSpace, 2016.
20. X. Yao et al., "Design and Fabrication of Robotic Arm for Multiple Degrees of Freedom Control," Dean Francis Press, 2023.
21. T. Baltrusaitis, A. Zadeh, Y. C. Lim and L. -P. Morency, "OpenFace 2.0: Facial Behavior Analysis Toolkit," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018, pp. 59-66, doi: 10.1109/FG.2018.00019.
22. Zhongmin Liu, Yuxi Peng, Wenjin Hu, Driver fatigue detection based on deeply-learned facial expression representation, *Journal of Visual Communication and Image Representation*, Volume 71, 2020, 102723, ISSN 1047-3203, <https://doi.org/10.1016/j.jvcir.2019.102723>.
23. J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
24. P. Ekman and E. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*, 2nd ed. Oxford, U.K.: Oxford Univ. Press, 2021.
25. Roboflow, "Facial Expressions Detection YOLOv5." [En línea]. Disponible en: <https://universe.roboflow.com/nicks-workplace/facial-expressions-detection-yolov5>. [Accedido: 06-dic-2025].

Estimulación Cognitiva mediante Juegos Serios: Aceptación de la Televisión Digital y Análisis Comparativo de la Respuesta Emocional Multidispositivo en Adultos Mayores

Ana Camacho¹[0000-0001-6689-8667], María José Abasolo¹²[0000-0003-4441-3264]
Rita Oliveira³[0000-0001-6041-9469], Mirian Solis⁴[0009-0004-0747-4185]
Galo Cornejo⁵[0000-0003-3523-2795]

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 Comisión de Investigaciones de la Provincia de Buenos Aires (CICPBA)-Facultad de Informática de la Universidad Nacional de la Plata (UNLP), Argentina

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

⁴ Facultad de Psicología, Educación y Comunicación-Universidad Católica de Santiago de Guayaquil (UCSG), Ecuador

⁵ Facultad de Ingeniería-Universidad Católica de Santiago de Guayaquil (UCSG), Ecuador

ana.camachoc@info.unlp.edu.ar, mjabasolo@lidi.info.edu.ar, ritaoliveira@ua.pt,
mirian.solis@cu.ucsg.edu.ec, galo.cornejo@cu.ucsg.edu.ec

Resumen. Cogni-TV es una plataforma de juegos serios diseñada para la estimulación cognitiva de adultos mayores en etapa preclínica de Alzheimer, cuya eficacia y aceptación han sido evaluadas mediante un estudio longitudinal de 24 semanas con 18 participantes. La investigación analiza la experiencia del usuario en tres interfaces digitales: televisión, tableta y teléfono móvil, empleando los modelos TAM (Aceptación Tecnológica) y SAM (Respuesta Emocional), junto con el test MoCA para el desempeño cognitivo. En este trabajo, el análisis de la respuesta emocional (SAM) se aborda de forma comparativa entre los tres dispositivos mencionados; sin embargo, en esta publicación solo presentamos el TAM de la interfaz televisiva comparado con el del estudio piloto previo, realizado con una muestra menor, de 5 participantes. Los hallazgos demuestran que los niveles de aceptación para la televisión se mantuvieron en valores altos y consistentes, validando la robustez de la plataforma a pesar del incremento en el tamaño de la muestra y el tiempo de exposición. Asimismo, el test SAM confirma que la televisión genera una experiencia emocional más favorable y con mayor percepción de control frente a los dispositivos móviles. Finalmente, los resultados del MoCA sugieren una estabilidad de las funciones ejecutivas tras la intervención. Se concluye que la televisión digital interactiva se posiciona como el medio más accesible y familiar para la tele-rehabilitación cognitiva en el hogar.

Keywords: Televisión Digital Interactiva, Alzheimer preclínico, Adultos Mayores, Aceptación de la Tecnología, Usabilidad, Juegos Serios, Memoria

1 Introducción

El envejecimiento progresivo de la población y el incremento en la prevalencia de deterioro cognitivo leve y enfermedad de Alzheimer en fases iniciales representan un desafío crítico para los sistemas de salud, especialmente en contextos latinoamericanos con recursos limitados. Esta problemática es evidenciada por autores como Rivas-Sucari y Rodríguez-Eguizabal, quienes sostienen que la salud cognitiva debe consolidarse como una prioridad en la agenda de salud pública regional, dado que el impacto socioeconómico de la dependencia funcional resulta potencialmente insostenible para los modelos de protección social vigentes. [1]. En etapas preclínicas, cuando las personas aún conservan su autonomía, las intervenciones no farmacológicas de estimulación cognitiva resultan fundamentales para preservar la funcionalidad y la calidad de vida. Al respecto, en el artículo de autoría [2] se destaca que las Tecnologías de la Información y Comunicación (TIC) orientadas a este perfil de usuario no solo facilitan el acceso a dichas intervenciones, sino que permiten un monitoreo continuo de la evolución cognitiva, actuando como un puente esencial entre el entorno clínico y la vida cotidiana del paciente.

Si bien las tecnologías digitales han emergido como soporte para programas de entrenamiento cognitivo, gran parte de estas soluciones se basan en computadoras, tabletas o teléfonos inteligentes, dispositivos que frecuentemente presentan barreras de accesibilidad para los adultos mayores. Factores como la brecha digital, la disminución de la precisión motriz y la alta carga visual limitan su efectividad. En contraste, la televisión, al ser un dispositivo cotidiano y familiar, ofrece un canal prometedor para implementar la Televisión Digital Interactiva (TVDi) como plataforma de intervención cognitiva. Al respecto, Webster y Valverde señalan que la falta de una alfabetización digital adecuada en los adultos mayores profundiza la exclusión tecnológica, especialmente cuando se enfrentan a interfaces complejas o dispositivos móviles [3].

En este contexto se desarrolló Cogni-TV, una plataforma basada en TVDi mediante tecnología WebView que integra tres juegos serios orientados a la memoria y la atención. Se postula que una interfaz televisiva, por su accesibilidad y simplicidad, puede optimizar el desempeño cognitivo, la usabilidad percibida y la experiencia emocional en comparación con dispositivos táctiles.

La viabilidad de esta propuesta se sustenta en un estudio piloto previo que evaluó de forma transversal la interacción con la plataforma Cogni-TV [4,5]. El presente artículo profundiza en dichos hallazgos mediante un estudio longitudinal de 24 semanas con 18 adultos mayores en etapa preclínica de Alzheimer. En esta publicación, se analiza de manera comparativa la respuesta emocional (SAM) entre televisión, tableta y teléfono móvil tras completar un ciclo de 12 sesiones por dispositivo. Asimismo, se presenta una revalidación del modelo de aceptación tecnológica (TAM) focalizada exclusivamente en la interfaz televisiva, contrastando los resultados actuales con la muestra inicial del estudio piloto. Para determinar la viabilidad integral de la propuesta, la intervención incorporó métricas de desempeño en los juegos y una valoración clínica mediante el test MoCA, permitiendo validar la robustez de la televisión digital interactiva como el canal predilecto para la tele-rehabilitación cognitiva en el hogar.

Este enfoque integral permite obtener evidencia sobre la viabilidad de la TVDi para la telerehabilitación cognitiva. El resto del artículo se organiza de la siguiente manera: la Sección 2 presenta la plataforma Cogni-TV; la Sección 3 describe el diseño metodológico del estudio longitudinal; la Sección 4 detalla los resultados de los análisis cuantitativos y cualitativos; y, finalmente, la Sección 5 expone las conclusiones, limitaciones y líneas futuras de investigación.

2 Programa de mantenimiento cognitivo a través de juegos serios

2.1 Infraestructura tecnológica

La plataforma **Cogni-TV** se sustenta en una arquitectura flexible diseñada para operar de forma multiplataforma en televisores, tabletas y teléfonos móviles. Para la interfaz televisiva, se desarrolló una aplicación nativa para Android TV basada en tecnología *WebView*, lo que permite la ejecución de los juegos terapéuticos directamente en pantalla sin requerir descargas previas [6]. El contenido se distribuye mediante protocolos web que garantizan una carga dinámica y estable de los recursos interactivos.

El repositorio ha sido desarrollado íntegramente con estándares web (HTML5, PHP, JavaScript y CSS), asegurando la interoperabilidad y una integración fluida con la plataforma de ejecución. Asimismo, la interfaz fue optimizada específicamente para la navegación mediante control remoto y la visualización en grandes dimensiones, garantizando que el diseño centrado en el usuario responda a las necesidades de la población objetivo.

Como dispositivo principal de pruebas, se utilizó un televisor DIGIO de 32 pulgadas (modelo DIG32HLC01). Posteriormente, se adaptó la aplicación para tabletas y teléfonos móviles. Se emplearon tabletas XTRATECH (modelo X8MT87 de 8 pulgadas) y teléfonos móviles Samsung Galaxy A02 de 6 pulgadas (ver Fig. 1).

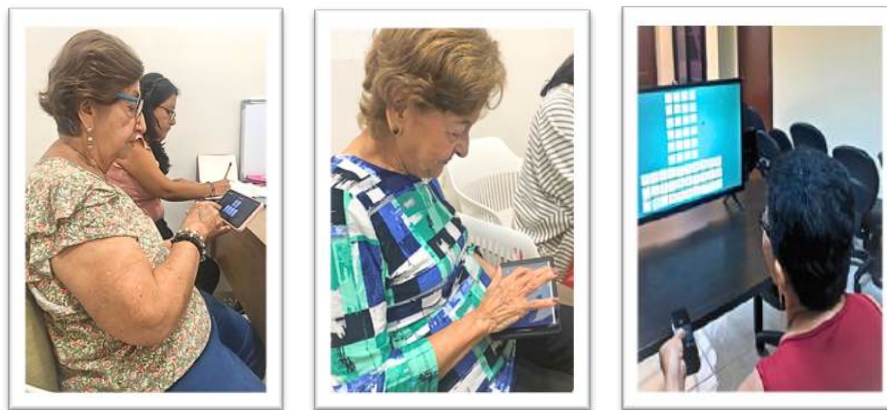


Fig.1. Uso de la Plataforma Cogni-TV en Múltiples Dispositivos

2.2 Repositorio de juegos serios para la intervención

El programa de intervención incluye tres juegos serios orientados a procesos de memoria y atención [7]: Banderas (reconocimiento visual y memoria geográfica), Cartas (memoria de trabajo visual) y Palabra Misteriosa (procesamiento del lenguaje y formación de palabras). Cada juego cuenta con tres niveles de complejidad progresiva (bajo, medio y alto).

La validación de Cogni-TV se sustentó en una colaboración interdisciplinaria entre investigadores del área de Ciencias de la Computación de la Universidad Católica de Santiago de Guayaquil y especialistas en neuropsicología del Centro Gerontológico 'Dr. Arsenio de la Torre Marcillo'. Esta alianza fue determinante para el desarrollo de la plataforma, cuyos fundamentos de diseño centrado en el usuario y usabilidad fueron validados inicialmente mediante una evaluación heurística [8]. Dicho estudio previo permitió refinar la interfaz y asegurar que las actividades interactuadas estuvieran técnicamente optimizadas y ajustadas a las capacidades cognitivas de adultos mayores en etapa preclínica de Alzheimer.

3 Metodología del Estudio Longitudinal de Cogni-TV

3.1 Diseño Experimental y Participantes

La investigación se basó en un diseño longitudinal cuasiexperimental con mediciones pre-test y post-test. Este diseño permite evaluar el impacto de herramientas digitales en contextos reales donde no es posible la aleatorización total, garantizando la validez de los resultados a través del seguimiento sistemático de los sujetos. El presente estudio expande y profundiza los alcances del piloto previo, en el cual se inició una intervención con 5 participantes femeninas centrada en la interfaz televisiva (8 sesiones) [4,5]. Durante dicha fase inicial, surgió la necesidad de contrastar la experiencia de uso con otros terminales como tableta y teléfono móvil, lo que derivó en una evaluación comparativa preliminar mediante los modelos **TAM** y **SUS**.

A partir de esos hallazgos, la investigación actual se planificó desde su origen como un estudio comparativo integral. En esta etapa, una cohorte de 18 participantes completó un ciclo de 12 sesiones por cada dispositivo (televisión, tableta y teléfono móvil). Esta progresión metodológica ha sido fundamental para determinar la viabilidad técnica de la plataforma Cogni-TV e identificar el impacto diferenciado de cada interfaz en la respuesta emocional y la adherencia al tratamiento. La **Tabla 1** muestra la comparativa detallada entre ambos estudios, omitiendo en esta publicación los resultados de la estimulación analógica (papel) para centrar el análisis en la interacción digital.

La muestra estuvo constituida por adultos mayores residentes del Centro Gerontológico "Dr. Arsenio de la Torre Marcillo". El proceso contó con el acompañamiento de un equipo interdisciplinario integrado por la investigadora principal, especialistas en neuropsicología y personal técnico-pedagógico. Se aplicaron criterios de elegibilidad estrictos: los de inclusión comprendieron sujetos de entre 65 y 85 años con diagnóstico de riesgo temprano o etapa preclínica de Alzheimer —operacionalizado mediante el test MoCA— y agudeza visual suficiente para la interacción digital. Los criterios de exclusión se centraron en la presencia de deterioro

cognitivo severo o limitaciones psicomotrices agudas que imposibilitaran el uso de los periféricos.

Bajo estas directrices, se conformó un grupo experimental de 18 participantes que utilizó la plataforma Cogni-TV. El perfil tecnológico inicial de la muestra reveló un uso diario de televisión mediante control remoto (100%) y un uso esporádico de smartphones para mensajería (30%), sin experiencia previa en tabletas. Esta familiaridad de base con el ecosistema televisivo es un factor crítico, ya que minimiza la carga cognitiva inicial y diferencia la curva de aprendizaje frente a dispositivos táctiles.

Tabla 1. Comparación Metodológica entre el Estudio Piloto 2024 y el Estudio Ampliado 2025.

Aspecto Comparado	Estudio Piloto 2024	Estudio Ampliado 2025
Tamaño muestral	5 participantes	18 participantes
Tipo de estudio	Exploratorio-descriptivo	Longitudinal, comparativo y con análisis multivariable.
Dispositivos evaluados	Se planificó sólo para televisión. Al finalizar agregamos comparativa con tableta y teléfono móvil.	Planificado como comparativa entre televisión, tableta y teléfono móvil.
Protocolo	Duración total: 10 semanas con 2 sesiones semanales de 20 minutos por sesión -8 sesiones con televisión -4 sesiones con tableta -4 sesiones con teléfono móvil	Duración total: 24 semanas con 1-2 sesiones semanales de 30 minutos por sesión -12 sesiones con televisión -12 sesiones con tableta -12 sesiones con teléfono móvil
Herramientas de análisis	Excel (promedios manuales)	SPSS con tablas cruzadas, distribuciones y análisis por dispositivo.
Instrumentos aplicados	Test TAM reducido y test SUS	Test TAM y test SAM
Datos cualitativos	Entrevistas breves y observación informal.	Observación estructurada más entrevistas a las participantes.
Evaluación cognitiva	MiniCog (pre-post)	MoCA (pre-post)

3.2 Instrumentos

Para evaluar la aceptación, utilidad y experiencia emocional se emplearon instrumentos complementarios:

- **Modelo de Aceptación Tecnológica (TAM):** Evaluó la utilidad percibida, facilidad de uso y la intención de uso [9].
- **Self-Assessment Manikin (SAM):** Midió la respuesta emocional mediante pictogramas de valencia y activación, facilitando la interpretación en adultos mayores [10].
- **Test MoCA (Montreal Cognitive Assessment):** Administrado por especialistas en neuropsicología al inicio y al final del estudio para valorar cambios en el desempeño cognitivo [11].

Los datos obtenidos mediante el TAM y SAM fueron procesados estadísticamente en SPSS para garantizar su consistencia y fiabilidad.

3.3 Protocolo de Intervención y Validación

La validación de la propuesta incluyó una revisión técnica y clínica que garantizó la pertinencia y seguridad de Cogni-TV antes de su implementación. Las neuropsicólogas evaluaron la adecuación cognitiva, carga mental y coherencia terapéutica, asegurando que la plataforma cumpliera criterios técnicos y clínicos para su uso seguro en adultos mayores en etapa preclínica de Alzheimer. Esta rigurosidad se alinea con Ro et al. (2023), quienes demuestran que los juegos interactivos mejoran la función cognitiva siempre que la complejidad de las tareas esté equilibrada para evitar la frustración [12].

El grupo experimental completó un programa extendido de 24 semanas, con una frecuencia de entre una y dos sesiones semanales de 30 minutos. Esta intervención se distribuyó de manera equitativa para evaluar la versatilidad de la plataforma: se realizaron 12 sesiones en televisión, 12 en tableta, 12 en teléfono móvil. Esta intervención se sustenta tanto en las recomendaciones dadas por el equipo de neuropsicólogos y psicólogos de gerontológico así como en los hallazgos en investigaciones tal como Lim et al. [13], quienes demostraron que los programas de entrenamiento basados en juegos serios son eficaces para mejorar la función cognitiva global y la memoria en adultos mayores con deterioro cognitivo leve, siempre que mantengan una regularidad y una duración prolongada para consolidar los beneficios terapéuticos.

La efectividad se evaluó al inicio y al final mediante el test MoCA, mientras que la aceptación tecnológica (TAM) y la experiencia emocional (SAM) se aplicaron al finalizar cada modalidad tecnológica.

4 Resultados: Análisis de la Respuesta Emocional y Revalidación de la Aceptación Tecnológica y Discusión

En esta sección se presentan los hallazgos derivados del estudio longitudinal de 24 semanas. El análisis se divide en tres ejes: la respuesta emocional ante los diferentes dispositivos (SAM), la revalidación de la aceptación tecnológica en la plataforma principal (TAM-TV) y la estabilidad del perfil cognitivo (MoCA).

4.1. Análisis Comparativo de la Respuesta Emocional (SAM)

Para cuantificar la experiencia subjetiva de los 18 participantes, se aplicó el modelo *Self-Assessment Manikin* (SAM) tras las sesiones en cada interfaz. Como se observa en la Figura 2, los resultados revelan diferencias numéricas significativas en la percepción del usuario según el dispositivo utilizado:

- **Valencia (Agrado/Bienestar):** La interfaz televisiva alcanzó niveles de agrado superiores, con un promedio de 8/9, frente al 5/9 registrado para el teléfono móvil. Estos datos sugieren que el formato de pantalla grande actúa

como un factor protector que mitiga la frustración visual y promueve un estado de ánimo positivo durante la estimulación.

- **Activación (Alerta/Energía):** En esta dimensión, la televisión logró mantener un estado de alerta óptimo con un promedio de 7,5/9. Por el contrario, el uso del móvil promedió un 5,2/9, lo que indica que las interfaces pequeñas y la carga cognitiva de la interacción táctil pueden inducir fatiga prematura en el adulto mayor, reduciendo su capacidad de mantener la atención en el juego (ver Fig. 2).
- **Dominio (Control/Empoderamiento):** Esta dimensión presentó la brecha más crítica entre dispositivos. Los usuarios reportaron un control significativamente mayor con la televisión (8/9) en comparación con el teléfono móvil (que descendió a un 1,5/9). Este resultado revalida la hipótesis de que el control remoto es una interfaz mucho más empoderadora y natural para esta población que el control táctil, el cual genera una sensación de falta de autonomía.

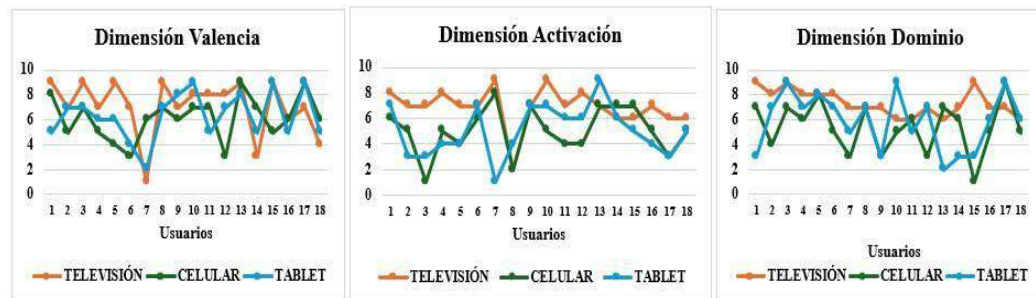


Fig. 2. Resultado global del Test SAM para cada dispositivo

4.2 Revalidación de la Aceptación Tecnológica (TAM) de la Interfaz Televisiva

Con el fin de verificar la robustez de la propuesta, se realizó una revalidación de los indicadores del modelo TAM para la interfaz de televisión, contrastando los datos actuales ($n=18$, 12 sesiones utilizando TVDi) con el estudio piloto de 2024 ($n=5$, 8 utilizando TVDi). Como se detalla a continuación y se visualiza en la Figura 3, los resultados revelan una transición desde una aceptación inicial entusiasta hacia una validación pragmática y robusta de la plataforma Cogni-TV:

- **Utilidad Percibida (PU):** Esta dimensión muestra el crecimiento más significativo. Mientras que en 2024 el acuerdo total sobre la utilidad se situaba en un 40%, en 2025 esta valoración escaló hasta el **77,78%**. Este incremento demuestra que la experiencia sostenida con los juegos de memoria permitió al usuario trascender la interfaz para valorar el sistema como una herramienta de salud legítima, duplicando la percepción del beneficio terapéutico tras completar las sesiones.
- **Facilidad de Uso Percibida (PEU):** En 2024, la percepción de facilidad fue alta (80% en "Totalmente de acuerdo"). Para 2025, con una muestra mayor y

más diversa, esta cifra se situó en el **72,22%**. Aunque aparece un 11,11% de usuarios en desacuerdo (atribuible a la mayor heterogeneidad de la muestra actual), la gran mayoría continúa validando que la adaptación de Cogni-TV para control remoto eliminó las barreras operativas detectadas en los dispositivos móviles (ver Fig. 3).

- **Actitud hacia el Uso (AU):** En el piloto de 2024, la actitud fue unánime (100% de acuerdo), impulsada por el "efecto novedad". En 2025, la actitud se vuelve más realista pero sigue siendo predominantemente positiva, superando el **83%** de aceptación combinada (entre "De acuerdo" y "Totalmente de acuerdo"). Esta transición hacia una actitud más diversificada es propia de procesos de adopción tecnológica a largo plazo tras superar la fase inicial de curiosidad.
- **Intención de Uso Futuro (ITU):** La televisión se confirma como el dispositivo con mayor proyección de uso recurrente. Mientras que en 2024 el 80% proyectaba un uso futuro total, en 2025 la intención de uso en la escala más alta se sitúa en el **61,11%**, con un **16,67%** adicional en "De acuerdo". La persistencia de esta intención mayoritaria tras 24 semanas de intervención confirma la viabilidad de Cogni-TV como herramienta de tele-rehabilitación, siendo la interfaz televisiva el factor determinante para asegurar la adherencia del adulto mayor.

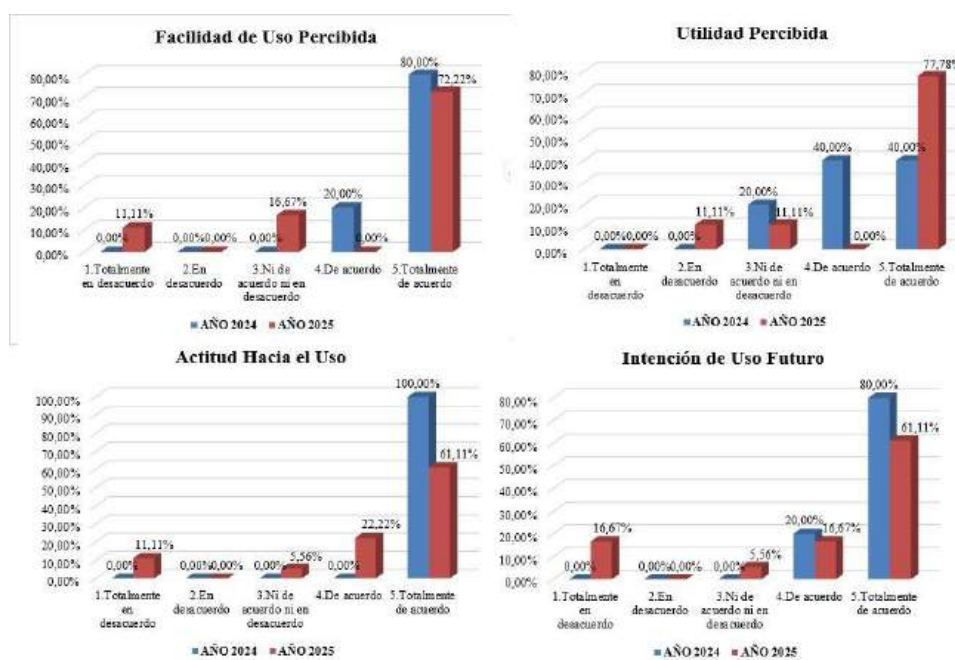


Figura 3. Resultados del test TAM sobre el uso de TVDi (estudio piloto de 2024 y estudio actual 2025)

4.3. Evaluación del Desempeño Cognitivo (MoCA)

El seguimiento clínico mediante el test MoCA, administrado por especialistas en neuropsicología al inicio y al cierre de las 24 semanas, arrojó resultados positivos en términos de mantenimiento:

- **Estabilidad de las funciones:** No se registraron descensos significativos en las puntuaciones globales de los 18 participantes.
- **Efecto de la intervención:** Dada la naturaleza degenerativa de la etapa preclínica de Alzheimer, la estabilidad de los resultados sugiere que el programa de estimulación cognitiva mediante Cogni-TV cumple su objetivo de preservar la funcionalidad ejecutiva y la memoria de trabajo durante el periodo evaluado.

5 Conclusiones

Este estudio longitudinal permitió revalidar la viabilidad de la plataforma **Cogni-TV** como herramienta de estimulación cognitiva en adultos mayores en etapa preclínica de Alzheimer. Al contrastar la experiencia de uso entre televisión, tableta y teléfono móvil, los hallazgos consolidan a la televisión como la interfaz más idónea para esta población debido a su familiaridad y ergonomía.

En primer lugar, los resultados del modelo SAM demuestran una clara superioridad de la televisión en términos de respuesta emocional. La brecha observada en la dimensión de dominio confirma que el control remoto ofrece una interacción mucho más empoderadora y natural que las interfaces táctiles. Estas últimas, especialmente en el formato de teléfono móvil, generaron mayores niveles de frustración y fatiga visual, lo que dificulta la adherencia a largo plazo.

En segundo lugar, la revalidación del modelo TAM para la interfaz televisiva permitió confirmar que la aceptación detectada en el estudio piloto previo se mantiene sólida tras un uso prolongado y con una muestra más amplia. El incremento en la Utilidad Percibida sugiere que, tras completar el ciclo de sesiones, los participantes logran trascender la novedad tecnológica para valorar el sistema como una herramienta legítima de salud. Aunque la actitud hacia el uso se volvió más pragmática y diversa con el tiempo, la intención de uso futuro sigue siendo mayoritaria, validando la robustez de la plataforma principal.

En el ámbito clínico, la estabilidad observada en los puntajes del test MoCA tras la intervención sugiere un efecto de mantenimiento funcional positivo. Dada la naturaleza degenerativa de la patología, este resultado es prometedor, indicando que la estimulación constante mediante juegos serios puede contribuir a preservar la autonomía y las funciones ejecutivas del adulto mayor.

Finalmente, este trabajo aporta evidencia sobre la eficacia de la Televisión Digital Interactiva como eje central para la tele-rehabilitación cognitiva en el hogar. La convergencia entre una respuesta emocional favorable y una aceptación tecnológica

consistente posiciona a Cogni-TV como una solución accesible, sentando las bases para futuras implementaciones a mayor escala que profundicen en el impacto de estas intervenciones en la calidad de vida de los pacientes.

Referencias

1. Rivas-Sucari, H. C., & Rodríguez-Eguizabal, J. L. (2024). Cognitive health in older adults. A public health challenge. *Gaceta medica de Mexico*, 160(2), 223-224.
2. Camacho, A., Abásolo Guerrero, M. J., & Oliveira, R. (2024). Las TIC orientadas a las personas mayores con demencia temprana de Alzheimer. In XI Ibero-American Conference on Applications and Usability of Interactive Digital Television (jAUTI 2022) (Córdoba, España, 17 y 18 de noviembre de 2022).
3. Arteaga Webster, M. I., & Tenecora Valverde, D. A. (2019). Alfabetización digital para adultos mayores (Bachelor's thesis, Universidad del Azuay).
4. Camacho, A., Abásolo, M.J., Oliveira, R. (2026). User Experience with Serious Games for Memory Stimulation in Older Adults: A Comparative Study Across Television, Tablet, and Smartphone Platforms. In: Abásolo, M.J., De Castro Lozano, C., Olmedo Cifuentes, G.F. (eds) Applications and Usability of Interactive TV. jAUTI 2024. Communications in Computer and Information Science, vol 2514. Springer, Cham. https://doi.org/10.1007/978-3-032-01162-6_5
5. Camacho, A., Abásolo Guerrero, M. J., & Oliveira, R. (2025). Experiencia de adultos mayores con juegos serios para estimular la memoria en tv interactiva: un análisis de facilidad de uso y utilidad percibida según el modelo TAM. In XIII Iberoamerican Conference on Applications and Usability of Interactive Digital TV (jAUTI)(República Dominicana , 13 al 15 de noviembre de 2024).
6. Tomalá, J.(2024). Diseño e implementación una plataforma de juegos terapéuticos para seguimiento y monitoreo de pacientes con Alzheimer en fase preclínica a través de la televisión digital interactiva con enfoque en dispositivos Android Tv (Trabajo de Integración curricular, Universidad Católica de Santiago de Guayaquil). <http://repositorio.ucsg.edu.ec/handle/3317/22776>
7. Camacho, A., Abásolo, M. J., & Oliveira, R. (2023). Hacia un portafolio de Juegos Serios para la Estimulación Cognitiva y de la Memoria de Personas con Alzheimer usando la Televisión Interactiva: Indagación en un Centro Gerontológico. In Proceedings of the XII Iberoamerican Conference on Applications and Usability of Interactive TV jAUTI.
8. Camacho, A., Abásolo, M. J., & Oliveira, R. (2024). Evaluación Heurística de Cogni-TV para Estimulación Cognitiva de Personas Mayores. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E75), 601-618.
9. Hueso, A. M., García, A. V. M., & Fincias, P. T. (2022). Revisión sistemática de aceptación de la tecnología digital en personas mayores. Perspectiva de los modelos TAM. *Revista Española de Geriatria y Gerontología*, 57(2), 105-117.
10. Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.
11. Holguín Zambrano, M., & Medina Quiroga, S. (2017). Diagnóstico de deterioro cognitivo leve demencia Senil Alzheimer mediante la utilización de MMSEMOCA (Bachelor's thesis). Recuperado de <https://acortar.link/2wnrX>.

12. Ro, D., Lee, J., Lee, G., Shin, S., & Kim, Y. H. (2023). Effect of interactive multitouch game-based cognitive intervention on cognitive function in older adults: A randomized controlled trial. *Digital Health*, 9, 20552076231176648.
13. Lim, E. H., Kim, D. S., Won, Y. H., Park, S. H., Seo, J. H., Ko, M. H., & Kim, G. W. (2023). Effects of home based serious game training (Brain Talk™) in the elderly with mild cognitive impairment: randomized, a single-blind, controlled trial. *Brain & NeuroRehabilitation*, 16(1), e4.



Framework liviano y abierto para el desarrollo de Presentaciones interactivas de Realidad Extendida

Martín Fernando Puricelli¹[0009–0004–5410–4611], María José Abásolo Guerrero²[1111–2222–3333–4444], and Roberto Ariel Guerrero³[0009–0009–7602–3253]

¹ Instituto de Tecnología e Ingeniería, Universidad Nacional de Hurlingham, Argentina

fernando.puricelli@unahur.edu.ar

² Facultad de Informática, Universidad Nacional de La Plata, Argentina

mjabasolo@lidi.info.unlp.edu.ar

³ Laboratorio de Computación Gráfica, Universidad Nacional de San Luis, Argentina

rag@unsl.edu.ar

Abstract. Este artículo presenta el diseño e implementación de un framework liviano y de código abierto para la creación y transmisión de presentaciones en Realidad Extendida (RE) mediante tecnologías web. La propuesta busca democratizar el acceso a experiencias inmersivas tridimensionales con fines educativos y de simulación, integrando hardware de bajo costo, interfaces amigables y funciones de visualización compartida en tiempo real. La arquitectura se basa en el framework ATON, mejorado con funcionalidades adicionales como control de puntos de vista (POV), renderizado 3D anaglifo e integración de RE mediante el uso de cámaras de dispositivos móviles.

Keywords: Realidad Mixta · Realidad Extendida · Visualización 3D · Escenas Virtuales · Dispositivos Móviles.

1 Introducción

El auge de las tecnologías inmersivas ha ampliado las posibilidades de visualización, interacción y comunicación en ámbitos educativos, industriales y culturales. En este marco, la Realidad Extendida (RE) un término amplio que abarca todas las tecnologías inmersivas que fusionan o amplían los entornos reales y virtuales, integra Realidad Virtual (RV), Realidad Aumentada (AR) y Realidad Mixta (RM) y se destaca por su capacidad de articular entornos digitales y físicos, permitiendo la interacción con contenidos tridimensionales de manera significativa. No obstante, muchos sistemas RE de alta gama dependen de hardware costoso y plataformas propietarias, lo que restringe su adopción en contextos educativos o con recursos limitados. La simulación, el entrenamiento y la educación [1] [2] constituyen algunos de los ámbitos más beneficiados por estos desarrollos, favorecidos además por la integración de dispositivos móviles de uso

cotidiano y el aprovechamiento de sensores incorporados como acelerómetros y giroscopios [3]. Actualmente existe una amplia oferta de dispositivos de RV y RM, generalmente asociados a aplicaciones específicas y con un fuerte énfasis en la industria del videojuego [4]. Este contexto refuerza la necesidad de integrar estas tecnologías en los procesos de enseñanza y aprendizaje, respaldada por numerosas experiencias documentadas en educación [5]. Además de la experiencia del estudiante, resulta clave considerar las oportunidades que estas herramientas brindan al docente, lo que exige entornos y frameworks que articulen adecuadamente dimensiones técnicas y pedagógicas [6]. El presente trabajo describe el diseño e implementación de un framework liviano y de código abierto para presentaciones en realidad extendida, basado en estándares web y bibliotecas libres. La propuesta se apoya en la plataforma ATON, extendiendo sus funcionalidades mediante nuevas capacidades.

2 Diseño y Metodología

2.1 Objetivos de diseño

El marco propuesto fue concebido con el objetivo de permitir experiencias de RE accesibles y escalables, orientadas a la enseñanza, la capacitación y la exploración colaborativa. Los principales criterios de diseño son: Accesibilidad, Interactividad, Soporte multiusuario, Simplicidad de implementación, Modularidad y apertura.

2.2 Arquitectura del sistema

El sistema propuesto se basa en una extensión de la plataforma ATON [10], un framework XR basado en la web originalmente desarrollado para la visualización de patrimonio cultural. Sobre esta base se incorporaron módulos adicionales orientados a soportar escenarios educativos y de presentación, ampliando sus capacidades de interacción, control de escenas y distribución de contenidos. La arquitectura resultante se estructura en un conjunto de componentes funcionales que operan de manera integrada dentro del entorno web. En particular, el sistema incorpora los siguientes módulos principales: (1) Motor de renderizado Web3D, responsable de la visualización y gestión del pipeline gráfico en el navegador; (2) Capa de integración WebXR, que habilita el acceso a dispositivos y modos de visualización inmersivos; (3) Editor y controlador de escenas, destinado a la configuración y gestión de los elementos que componen el entorno virtual; (4) Sistema de gestión de puntos de vista (Points of View, POV), que permite la definición y conmutación de cámaras preconfiguradas durante la navegación; (5) Módulo de streaming, encargado de la transmisión y sincronización de la experiencia de visualización; y (6) Capa de sincronización multiusuario, que posibilita la interacción coordinada de múltiples participantes dentro de la misma escena. La figura: 1 presenta un esquema general de la arquitectura del sistema y la relación entre los diferentes módulos que componen el framework.

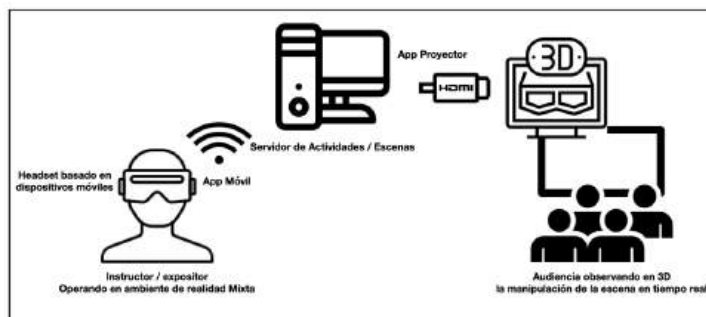


Fig. 1. Componentes de la Arquitectura

2.3 Modos de operación

El sistema admite tres modos principales, cada uno adaptado a diferentes casos de uso como se detallan en la tabla 1.

Table 1. Modos de uso de la plataforma

Modo	Funcionalidad	Caso de uso
Modo Editor	Permite al presentador construir y organizar la escena	Diseño de presentaciones educativas
Modo Aplicación	Proporciona una interfaz interactiva exploratoria	Exploración por estudiantes
Modo Transmisión	Sincroniza la vista de todos los usuarios	Presentaciones en aula o conferencias.

2.4 Roles de Usuario y Caracterización

El sistema está diseñado para contemplar tres roles principales: Presentador, Participante, Desarrollador/Técnico. Las actividades típicas incluyen carga de modelos, posicionamiento espacial, navegación por puntos de vista, anotaciones, control de transmisión en tiempo real e interacción. Basado en la taxonomía propuesta por Speicher [7], el framework implementado se clasifica según la tabla 2.

Table 2. Taxonomía de Speicher para el desarrollo

Dimensión	Características del framework
Número de entornos	Doble (real + virtual)
Número de usuarios	Multiusuario sincrónico
Nivel de inmersión	Moderado
Nivel de virtualidad	Equilibrado entre físicos y digitales
Grado de interacción	Alto (control de POV, anotaciones, manipulación)
Entrada/Salida	Táctil, teclado, sensores y cámara de móvil
Modalidad de control	Mixta (modos de usuario y control presentador)

2.5 Mejoras e intervenciones sobre la plataforma ATON

Con el fin de extender las capacidades originales del framework ATON para la gestión de escenas, la interacción del usuario y la visualización inmersiva en entornos Web3D, se implementó un conjunto de mejoras que requirieron modificaciones en el código base de ATON y la integración de módulos adicionales. Estas mejoras incluyen:

- la implementación de múltiples modos de acceso de usuario.
- el desarrollo de una interfaz de control que permite a los presentadores gestionar y cambiar entre puntos de vista predefinidos.
- la integración de renderizado anaglífico estereoscópico utilizando filtros Three.js.
- la adaptación del modo VR para admitir la visualización de realidad mixta utilizando cámaras móviles.

Estas intervenciones se resumen en la Tabla 3. El repositorio del código fuente se encuentra disponible en: <https://github.com/fpuricelli/zonar-v2>

Table 3. Módulos de ATON adaptados

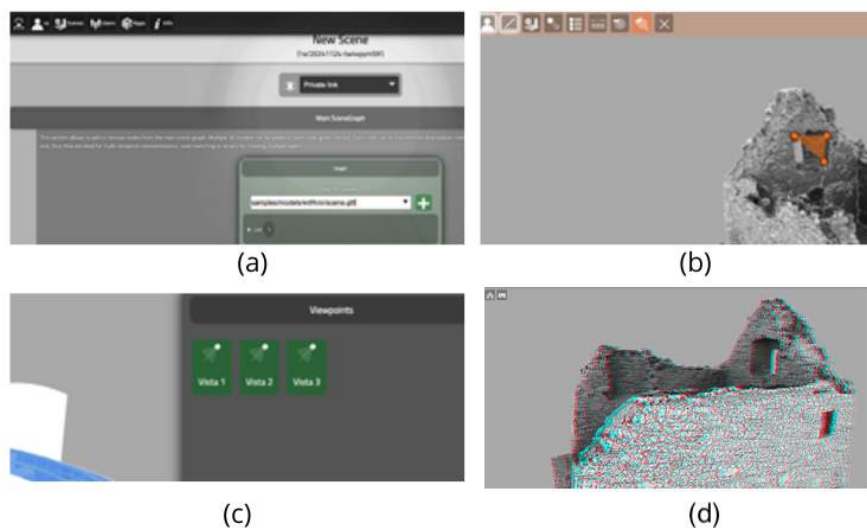
Módulo ATON	Adaptación implementada
Aton.XR.js	referencia a "immersive-vr"
ATON.mediaflow.js	generar la transmisión de cámara en vivo

Adicionalmente, se desplegó una versión demostrativa accesible en: <https://zonar-v2-docker.onrender.com/> Las principales extensiones implementadas se sintetizan en la Tabla 4.

Table 4. Código incorporado

Mejora implementada	Ruta - relativa a raíz del repositorio
Listar los POVs disponibles	/control/povs.html
Capturar parámetros del usuario y de la escena, y generar enlaces dinámicos para acceder a la diferentes vistas	/control/index.html
Acceder a la vista de Realidad Aumentada	/ag3d/
Incorporar el filtro Anaglyph de Three.js	/public/src/ATON.ag.js

La Figura 2 presenta el flujo de preparación y uso de escenas en la plataforma, incluyendo: (a) selección de recursos, (b) incorporación de anotaciones sobre objetos, manipulación de POVs durante la navegación y (d) visualización mediante 3D anaglifo.

**Fig. 2.** capturas de pantallas de diferentes funcionalidades

3 Implementación y casos de uso

El ecosistema XR actual dispone de una amplia variedad de componentes de software que facilitan la explotación de dispositivos inmersivos, incluyendo motores 3D [8] [9], SDKs especializados y bibliotecas orientadas a la ejecución en navegadores web, como WebXR. En este contexto, el framework propuesto se

implementó como una aplicación web nativa, ejecutable en navegadores modernos, basada en HTML5, WebGL y bibliotecas JavaScript. La arquitectura desarrollada permite su ejecución en dispositivos de escritorio y móviles, garantizando portabilidad y accesibilidad sin requerir instalaciones adicionales. Todos los componentes fueron evaluados en diferentes configuraciones de hardware y navegadores, verificando su funcionamiento en entornos heterogéneos. Actualmente, el sistema se encuentra en fase de evaluación de usabilidad, a partir de la implementación de diversos escenarios de aplicación representativos. Entre los casos de uso considerados se incluyen: (i) visualización de modelos anatómicos en contextos de formación en ciencias de la salud, (ii) recorridos arquitectónicos interactivos para la exploración de espacios tridimensionales y (iii) exploración de modelos técnicos en cursos de ingeniería. En todos los escenarios se habilitan funcionalidades comunes tales como exploración interactiva del entorno 3D, transmisión guiada por el instructor mediante control de puntos de vista y visualización mediante realidad extendida (XR), lo que permite estructurar experiencias de aprendizaje inmersivas accesibles desde la web. Los resultados iniciales indican que el sistema permite realizar presentaciones interactivas eficaces con unos requisitos de hardware mínimos.

4 Conclusiones y futuros desarrollos

Este artículo presenta el diseño e implementación de un framework abierto basado en la web para la creación y distribución de presentaciones de realidad mixta (MR) orientadas a contextos educativos y de simulación. El sistema se plantea como una alternativa accesible, escalable y de bajo costo frente a plataformas XR propietarias, priorizando apertura tecnológica y facilidad de uso. El framework extiende la plataforma ATON mediante la incorporación de capacidades como renderizado 3D anaglifo, visualización MR basada en cámara, transmisión multiusuario y control de puntos de vista (Points of View, POVs). Estas extensiones permiten a usuarios con conocimientos técnicos mínimos crear, configurar y compartir experiencias interactivas 3D directamente desde la web. El sistema se encuentra actualmente en fase de evaluación en distintos escenarios experimentales con el objetivo de analizar sus capacidades funcionales y su aplicabilidad en contextos educativos. Los principales aportes de este trabajo incluyen: (i) el desarrollo de un prototipo funcional de herramienta web nativa para presentaciones inmersivas, (ii) la integración de múltiples modos de visualización bajo una interfaz unificada y (iii) la exploración de su aplicación en entornos educativos. Las líneas futuras de investigación contemplan la mejora de la fidelidad de la experiencia mediante mapeo espacial y seguimiento basado en marcadores, la optimización de la interfaz mediante edición drag-and-drop y bibliotecas de recursos, la ampliación del soporte de hardware mediante OpenXR, así como la incorporación de arquitecturas modulares basadas en complementos. Asimismo, se prevé la realización de pilotos a mayor escala en aulas y entornos de formación, con el objetivo de evaluar escalabilidad, adopción y sostenibilidad de uso. En conjunto, el trabajo contribuye a la democratización de la realidad exten-

didia (XR) al proponer una solución web abierta, flexible y de bajo umbral de adopción para docentes, investigadores y desarrolladores interesados en explorar marcos de presentación inmersivos basados en la web.

References

1. Maraggi, S.J.M.: Realidad virtual y realidad aumentada: Aplicación al entrenamiento en escenarios de alto riesgo. In: *XXIII Concurso de Trabajos Estudiantiles (EST 2020) – JAIIO 49* (2020). <http://sedici.unlp.edu.ar/handle/10915/115883>
2. Martirosov, S., Kopecek, P.: Virtual Reality and its Influence on Training and Education—Literature Review. In: Katalinic, B. (ed.) *DAAAM Proceedings*, vol. 1, pp. 708–717. DAAAM International, Vienna (2017). doi:10.2507/28th.daaam.proceedings.100
3. Bauer, J., Ebert, A.: Mobile Devices for Virtual Reality Interaction. A Survey of Techniques and Metaphors. In: *Virtual Realities: International Dagstuhl Seminar, Dagstuhl Castle, Germany, June 9–14, 2013, Revised Selected Papers*, pp. 91–107. Springer, Cham (2015). doi:10.1007/978-3-319-17043-5_6
4. VRcompare: VRcompare. <https://vr-compare.com/> (accessed Oct 31, 2021)
5. Alvarado, Y., Jofré, N., Rosas, M., Guerrero, R.: Aplicaciones de Realidad Virtual y Realidad Aumentada como soporte a la enseñanza del Dibujo Técnico. pp. 65–74 (2019)
6. Quint, F., Sebastian, K., Gorecky, D.: A Mixed-Reality Learning Environment (2015)
7. Speicher, M., Hall, B.D., Nebeling, M.: What is Mixed Reality? In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–15 (2019). doi:10.1145/3290605.3300767
8. Wani, S.S., Kamble, S.B., Raman, R.: A Study and Comparative Analysis of Development Tools for Virtual Reality Applications. In: *2020 Seventh International Conference on Information Technology Trends (ITT)*, pp. 212–217 (2020). doi:10.1109/ITT51279.2020.9320893
9. Abásolo Guerrero, M.J., Manresa Yee, C., Más Sansó, R., Vénere, M.: *Realidad virtual y realidad aumentada*. Editorial de la Universidad Nacional de La Plata (EDULP), La Plata (2011). doi:10.35537/10915/18399
10. ATON APIs: ATON APIs Documentation. <https://aton.ispc.cnr.it/apidoc/client/>

Evaluating the Usability and Effectiveness of an Interactive Platform for Stakeholder Engagement and Empowerment

Enrickson Varsori¹ [0000-0002-8875-9278], Mariana Pinho² [0000-0002-0173-444X] Dionísia Laranjeiro¹ [0000-0003-3347-7967], Daniel Crespo² [0000-0003-4402-5229], Asta Hiippala³, Anna Broberg³ and Ana I. Lillebø² [0000-0002-5228-0329]

¹ Research Centre on Didactics and Technology in the Education of Trainers, CIDTFF, University of Aveiro, Aveiro, Portugal varsori@ua.pt

² ECOMARE, Centre for Environmental and Marine Studies – CESAM, Department of Biology, University of Aveiro, Aveiro, Portugal

³ Maptionnaire by Mapita Oy, Helsinki, Finland

Abstract. Empowering local communities is essential for the co-production of effective and sustainable environmental solutions that promote the restoration of marine ecosystems. This study evaluates the interactive usability and professionals' perceptions of a platform designed to support the engagement of coastal communities in implementing innovative solutions. The platform, Maptionnaire, was configured and used to facilitate stakeholder engagement in the NUTS II Centro Region of Portugal. The evaluation involved 21 professionals from various fields related to sustainable environmental practices, who provided feedback based on UX metrics, including the System Usability Scale (SUS), Self-Assessment Manikin (SAM), and AttrakDiff scale. Post-test data were also collected through a semi-structured technique designed to explore perceptions of the platform's functionalities, using a semi-structured questionnaire and a think-aloud protocol. The results obtained from the test indicate that implementing interactive tools is both appropriate and feasible for use by a wide range of stakeholders and for promoting their engagement. Furthermore, the results provide a foundation for the further development of a functional platform for participatory environmental action.

Keywords: User Experience, Usability Test, Restoration of Marine Ecosystem, Participatory Environmental Solutions

1 Introduction

Addressing the complex challenges facing Europe's Ocean and coastal regions requires participatory approaches to promote engagement and the co-development of innovative solutions to restore and enhance the resilience of coastal ecosystems to climate change. Empowering local actors through dialogue, co-creation, and co-development of these solutions is key to fostering sustainable transformation in marine

and coastal areas. One tool used for stakeholder engagement is the digital platform Maptionnaire, which collects data through an interactive, map-based survey that is easily accessible via mobile phones and computers. The survey for the present study was designed to reflect the key environmental pressures previously identified in the region, including coastal erosion, biodiversity loss, seagrass meadow fragmentation, and marine litter. Participants are invited to assess the severity of each pressure using multiple criteria and to actively contribute by adding new points on the map and suggesting challenges and solutions.

Thus, the present study focuses on an initial round of testing to systematize the evaluation process and to understand how users interact with the platform in the context of environmental participation. More specifically, the study evaluates the platform's usability and perceived effectiveness by combining SUS, SAM, and AttrakDiff metrics with qualitative perceptions. The contribution of this study lies in applying a multidimensional evaluation to a participatory platform and in translating the findings into practical recommendations for improving its functionality.

2 Platform Features Available to Users

2.1 Visualize, Analyze, and Suggest Points of Interest on the Map

The platform¹ presents a set of functionalities², implemented in complementary components that structure the user interaction process and support different forms of stakeholder input throughout the survey (see Figure 1). These functionalities were designed to guide users from information access to spatial contribution and structured evaluation, thereby enabling a more active form of participation.

Figure 1A illustrates the initial stage of interaction, in which users access a list of proposed solutions and their associated activities, accompanied by explanatory text that helps them understand the context, objectives, and scope of each proposal (e.g., “Resilience to coastal erosion for people and nature”). Figure 1B shows the georeferencing functionality, which enables users to identify specific areas on the map where selected activities may be relevant, thereby enabling the spatial localization of challenges or potential interventions. Figure 1C depicts the ranking procedure, where users order the activities within a given solution according to their perceived priority, providing comparative insights into stakeholders’ preferences. Figure 1D presents the multicriteria assessment interface, which requires users to evaluate each activity across

¹ The Maptionnaire platform was used as the operational interface for the solutions presented in the project. The functionalities, visualizations, and graphs observed correspond to the state of the platform at the time of the usability tests and may differ from its configuration, appearance, or capabilities in the version currently available.

² The platform has a set of functionalities: 1) decision support for restoration actions, 2) environmental sustainability of restoration activities, 3) social sustainability of restoration activities, 4) economic sustainability of restoration activities, 5) mapping and georeferencing areas of interest, 6) promoting ocean and environmental literacy, 7) encouraging participation in activities, and 8) promoting participation in citizen science.

predefined criteria, generating structured feedback on environmental, social, and economic considerations.

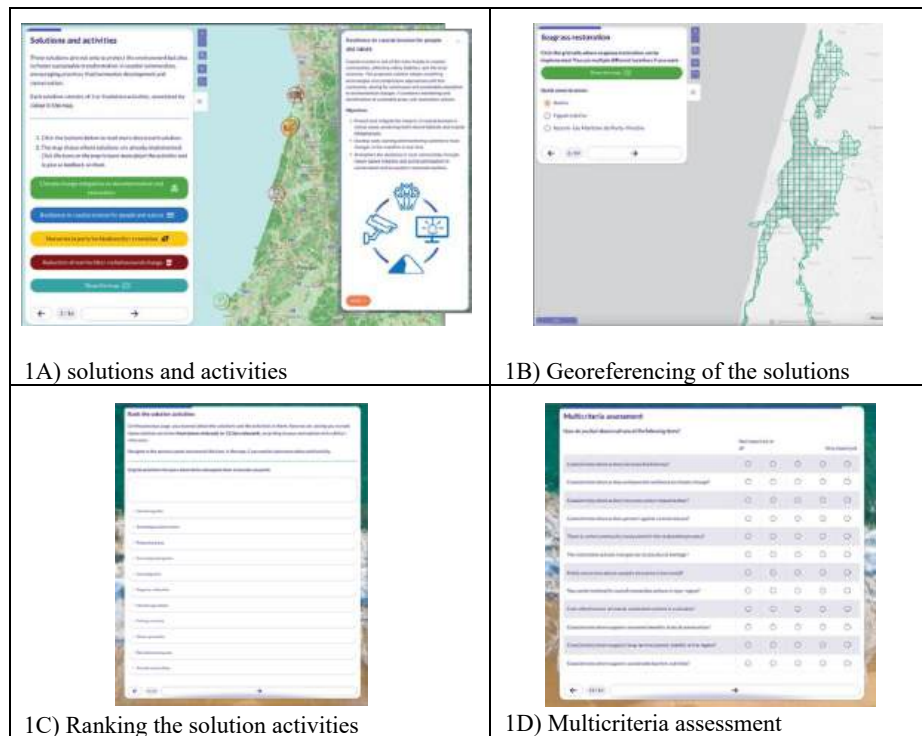


Fig. 1. Functionalities implemented in the platform.

3 Evaluation Methodology

The methodology is based on the principles of Human-Computer Interaction (HCI), with a focus on usability and user experience assessment in an initial laboratory-based evaluation of the platform. To evaluate the platform's features, a triangulation of usability tools was employed. Namely, SUS [1, 2], ATTRAIDIFF [3, 4], and SAM [5, 6], as demonstrated in other usability contexts [7]. The use of multiple instruments was intended to capture complementary dimensions of interaction: perceived usability, pragmatic and hedonic qualities, and emotional reactions during platform use. Additionally, a data collection technique developed by the research team, which focused on perceptions of platform functionalities, was used to collect professionals' perceptions and opinions.

The study followed a moderated post-use evaluation procedure. First, participants interacted with the platform and explored its main functionalities, including access to the proposed solutions and activities, georeferencing, ranking, and multicriteria assessment. Before the test session, the research team explained the platform's purpose

and the tasks to be completed. After using the platform, participants completed the selected usability and UX instruments and answered a set of questions to assess their perceptions of the platform's functionalities.

To evaluate the platform's usability in its current functional state, professionals from various fields were invited, all of whom were directly involved in sustainability-related areas relevant to the study context. This initial sample was selected to obtain informed feedback from participants who could critically reflect on the platform's relevance, clarity, and usefulness for environmental participation processes. The results were subsequently analyzed systematically. All data were evaluated according to the individual UX protocols of each scale, with the analysis prioritized according to the framework suggested by Almeida et al. [7], which assesses implementation characteristics in three key areas: instrumental qualities, non-instrumental qualities, and emotional reactions. Additionally, we followed Sauro's [8] guidelines for interpreting the SUS, which include scaling and evaluating it based on five main components: percentiles, grades, adjectives, acceptability, and, lastly, promoters and detractors.

Perceptions of the platform's functionalities. To assess professionals' perceptions of the platform, they were asked to rate their level of interest in the eight functions available in the system (see note 2). All items were evaluated on a 5-point scale, ranging from 1 (Very uninteresting) to 5 (Very interesting). They were also asked to indicate their level of agreement with the statement "The platform is suitable for engagement between different generations of users", using a 5-point scale from 1 (Strongly disagree) to 5 (Strongly agree). In addition, professionals identified at least two positive aspects of the platform, areas requiring improvement, and their three main motivations for using it. All responses were examined through thematic analysis, following a coding scheme developed for this study. Research team members independently performed open coding, after which the results were compared and discrepancies resolved.

Sample characterization. Data were collected from 21 professionals (8 male and 13 female) who participated in the Laboratory Test (InLab). The professionals' ages ranged from 25 to 60 years old ($M = 40.19$, $SD = 8.10$). The majority (62.5%) belonged to higher education institutions and included academics, non-faculty staff, and PhD students from various fields, including natural and social sciences, economics, and technology. Additionally, 12.5% belonged to public administration, 8.3% to business, and 8.3% to services. Participants were recruited through an invitation to attend a series of workshops organized within the scope of the project and focused on its thematic area. All data were anonymized, and only non-identifiable sociodemographic and professional information was used for characterization purposes, ensuring that no individual respondent could be identified.

4 Results and discussion

4.1 Interactive Evaluation Through UX Metrics

The results of the InLab tests are based on the triangulation of UX evaluation instruments, namely the SUS, AttrakDiff, and SAM scales. The findings are reported following previous studies [7], with additional descriptions regarding usability aspects in the SUS, as outlined by Sauro [8]. The total test scores are displayed in Table 1.

Table 1: Score According to the SUS, SAM and AttrakDiff Instruments.

Instrumental qualities	Non-instrumental qualities			Emotional reactions			
SUS (0 to 100)	AttrakDiff (-3 to 3)			SAM (1 to 5)			AttrakDiff (-3 to 3)
	QP	HQ-S	HQ-I	Sat.	Mot.	Cont.	ATT
75.6	1.06	1.33	0.44	4	4	3.8	1.71
UX Dimensions	Stimulation		Identification	Emotion		Aesthetics	

In terms of instrumental qualities, the SUS analysis indicates an overall performance of 75.6 ($SD = 11.1$; ranging from 55 to 95), with a general user average of 68, which represents a superior classification. The system is rated B on the scale, with a good indication regarding usability and the tools provided. The system is classified as Acceptable, meaning that its usability is within the parameters accepted by users. The Net Promoter Score classification places the professionals as Passive, indicating that, although the users consider the system good, they are not classified as Promoters. The SAM results indicate that professionals experienced a predominantly positive emotional response while using the platform. Motivation received the highest rating ($M = 4.05$; $SD = 0.6$), followed by satisfaction ($M = 3.95$; $SD = 0.6$) and sense of control ($M = 3.81$; $SD = 1.1$). Despite these positive outcomes, the higher standard deviation in the control dimension suggests some variability in professionals' experiences, suggesting that some users may have felt less confident navigating the platform. The AttrakDiff scale results were analyzed across four main dimensions: Perceived Quality (PQ), Attractiveness (ATT), Hedonic Quality (HQ), and Hedonic Experience (HE). Regarding structural dimensions, the PQ dimension received a moderate score of 1.06, consistent with the SUS results. This indicates that the system's usability is functional, and users do not face significant difficulties when interacting with it. However, the ATT dimension, related to the platform's aesthetic aspects, received the highest score, averaging 1.71. This result supports the qualitative feedback gathered from the perceptions of the platform's functionality responses (discussed in the next section), which highlighted the platform's aesthetics as a positive aspect. Specifically, the word pairs “ugly-attractive” and “bad-good” were most prominent in this dimension, reflecting the platform's pleasant visual impression.

Regarding the non-instrumental dimensions, the platform's scores varied across them. For the HQ-I dimension, the score was 0.44, indicating that users did not have an immediate identification with the interface. In contrast, the HQ-S dimension received

a score of 1.33, showing positive results in terms of originality, interest, and the encouragement generated by the system. Upon reviewing the word pairs, in the HQ-S dimension, the pair “dull-captivating” received the highest score, while in the HQ-I dimension, the pair “cheap-premium” received the lowest score.

In terms of the platform's evaluation by professionals, the dimensions Pragmatic Quality (QP = 1.06) and Hedonic Quality (QH = 0.89) received scores that indicate adequate pragmatic features, though there is still room for improvement to enhance the overall experience. The moderate score for QP suggests that the platform is functionally usable, but there is scope for refinement to improve ease and efficiency of task completion. Additionally, the portfolio representation from QP scores is positioned near the boundary between the “task-oriented” and “desired” areas. This positioning suggests that while the platform is still predominantly task-oriented, it is beginning to display features that make it more appealing to users, moving towards a more desirable experience. This shift towards the “desired” area is a positive indication, reflecting the platform's potential to offer a more engaging and stimulating experience, beyond its basic functionality.

Overall, the results indicate that the platform achieved a positive balance between usability and user experience in this initial evaluation. The SUS score suggests good usability, while the SAM and AttrakDiff results show that the platform was generally perceived as motivating, satisfactory, and attractive. These findings suggest that the platform is not only functionally adequate but also capable of supporting a positive interaction experience in participatory environmental contexts. At the same time, the lower score in the identity dimension and the qualitative comments regarding interface simplification, user control, and feedback mechanisms indicate that further refinements are still needed to improve the overall interaction flow and strengthen users' connection with the platform.

4.2 Perceptions of the Platform's features

Professionals' perceptions of the platform's functionalities were analyzed. The results show that, overall, professionals found all the platform's functions interesting. They found it most interesting to promote ocean and environmental literacy in coastal areas ($M = 4.29$, $SD = 0.72$) and to encourage participation in citizen science activities ($M = 4.29$, $SD = 0.96$). The component of promoting their participation in activities through the platform was rated as the least interesting, with an average rating of neither interesting nor uninteresting ($M = 3.09$; $SD = 0.99$). They also considered the platform suitable for engaging users across different generations ($M = 3.86$; $SD = 0.96$). Professionals ranked the assessment of restoration activities for economic sustainability as the least motivating factor for using the platform. From the analysis of the written answers, three key positive aspects emerged as central themes: 1) interface aesthetics; 2) user friendliness, and 3) platform features. When assessing areas for improvement, the three main aspects identified by professionals were simplifying the user interface, adding a save button, and improving feedback to users on the activities conducted.

These findings complement the UX metrics by showing that professionals not only evaluated the platform positively in quantitative terms but also recognized concrete

strengths in its visual appeal and overall ease of use. At the same time, the suggested improvements underscore the need to refine interaction clarity and feedback mechanisms, which are important for strengthening the platform's usability and perceived value in participatory environmental contexts.

5 Conclusions

The test results indicate that the platform demonstrated good usability ($SUS = 75.6$), with users also reporting high levels of satisfaction and motivation in the SAM assessment, with mean scores close to 4 on a 1-to-5 scale. The Pragmatic Quality ($PQ = 1.06$) and Hedonic Quality–Stimulation ($HQ-S = 1.33$) dimensions suggest an effective and emotionally engaging interaction, while Hedonic Quality–Identity ($HQ-I = 0.44$) reflects a more moderate level of identification with the platform. The global attractiveness score ($ATT = 1.71$) indicates positive professional acceptance. Overall, the triangulation of usability and user experience metrics yielded favorable results, showing that the platform is functionally adequate and positively perceived in the context of participatory environmental engagement.

These findings support the platform's potential to advance participatory environmental decision-making by providing a usable, engaging tool for collecting stakeholder input. At the same time, the evaluation identified areas for improvement, particularly in interface simplification, user control, and feedback mechanisms. Based on these results, adjustments are being made to improve usability and strengthen stakeholder engagement toward empowerment. Future work will focus on implementing these improvements and testing the platform with broader stakeholder groups in real participatory settings.

6 Limitations

The reliance on self-report measures is a limitation, as single-source self-reports may be influenced by social desirability and are therefore less reliable than multiple data sources. The study also relied on a small convenience sample, limiting the conclusions that can be drawn from the results. Therefore, findings cannot be used to establish the experiences with the platform in the larger population.

Acknowledgments. The work was funded by the A-AAGORA project (DOI: 10.3030/101093956) co-funded by the European Union under the Horizon Europe innovation action programme (Grant agreement ID: 101093956) in collaboration with RESTORE4Cs project (DOI: 10.3030/101056782) co-funded by the European Union under the Horizon Europe innovation action programme (Grant agreement ID: 101056782), and BESIDE project funded by the ERA Chair BESIDE project (DOI: 10.3030/951389) financed by the European Union's Horizon 2020 (Grant Agreement ID: 951389). To the professionals who took their time to share their knowledge and expertise with us. Their valuable contribution is much appreciated. The authors acknowledge financial support from CESAM by FCT (UIDP/50017/2020 + UIDB/50017/2020 + LA/P/0094/2020) and CIDTFF by FCT (UIDB/00194/2020), through national funds.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Brooke, J.: SUS: A “Quick and Dirty” Usability Scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., and Weerdmeester, B. (eds.) *Usability Evaluation In Industry*. pp. 189–194. CRC Press, London (1996). <https://doi.org/10.1201/9781498710411>.
2. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*. 24, 574–594 (2008). <https://doi.org/10.1080/10447310802205776>.
3. Hassenzahl, M., Diefenbach, S., Göritz, A.: Needs, affect, and interactive products - Facets of user experience. *Interacting with Computers*. 22, 353–362 (2010). <https://doi.org/10.1016/j.intcom.2010.04.002>.
4. Hassenzahl, M., Tractinsky, N.: User experience - A research agenda. *Behaviour and Information Technology*. 25, 91–97 (2006). <https://doi.org/10.1080/01449290500330331>.
5. Bradley, M.M., Lang, P.J.: Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*. 25, 49–59 (1994). [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9).
6. Morris, J.D.: Observations: SAM: The Self-Assessment Manikin; An Efficient Cross-Cultural Measurement of Emotional Response. *Journal of Advertising Research*. 35, 63–68 (1995).
7. Almeida, P., Abreu, J., Silva, T., Varsori, E., Oliveira, E., Velhinho, A., Fernandes, S., Guedes, R., Oliveira, D.: Iterative User Experience evaluation of a User Interface for the unification of TV contents. *Communications in Computer and Information Science*. 813, 44–57 (2018). https://doi.org/10.1007/978-3-319-90170-1_4.
8. Sauro, J.: 5 ways to interpret a SUS score, <https://measuringu.com/interpret-sus-score/>.

Enhancing Accessibility and Usability in DanceMove: Expert Evaluation of a Foot-Interaction Game for Older Adults

Ceci Diehl¹[0000-0002-0728-7492], Telmo Silva¹[0000-0001-9383-7659], Jorge Ferraz¹[0000-0002-0492-2307], Gonçalo Santinha²[0000-0002-4732-5959], Óscar Ribeiro³[0000-0003-4740-7951], Ana Martins⁴[0000-0002-5220-0528], Nelson Rocha⁵[0000-0003-3801-7249] and Anabela G. Silva⁶[0000-0002-4386-5851]

¹ DigiMedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal

² GOVCOPP, Department of Social, Political and Territorial Sciences, University of Aveiro, Aveiro, Portugal

³ RISE-Health, Department of Education and Psychology, University of Aveiro, Aveiro, Portugal

⁴ RISE-Health, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal

⁵ Department of Medical Sciences, Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Aveiro, Portugal

⁶ RISE-Health, School of Health Sciences, University of Aveiro, Aveiro, Portugal
cecidiehl@ua.pt, tsilva@ua.pt, jfa@ua.pt, g.santinha@ua.pt,
oribeiro@ua.pt, anaisabelmartins@live.ua.pt, npr@ua.pt and
asilva@ua.pt.

Abstract. Foot movements as a form of interaction with digital systems have been increasingly explored as an alternative input method in various technological solutions, particularly in video games. By using pressure-sensitive foot platforms, these technologies encourage physical activity, promote movement and well-being, and, through the playful and social aspects of gaming, enhance user engagement. Despite their recognised benefits, most modern games are primarily designed for young and adult audiences. When considering specific user groups, such as older adults, it is essential to adapt technologies and interfaces to their needs and characteristics. Within this context, DanceMove (SHAPES, Horizon 2020 Grant Agreement – 857159) was developed as a dance game for older adults, integrating choreographies, adjustable difficulty levels, and music selection to improve accessibility and user involvement. This study presents and discusses the results of a user interface evaluation of DanceMove, conducted with experts in design and communication sciences, who assessed the game considering both their professional knowledge and the characteristics of the target audience (seniors). Results from the System Usability Scale (SUS) indicated good usability (mean score = 75, SD = 14.2), with most experts agreeing that the game was easy to play, well integrated, and simple to learn. The Standardized User Experience Percentile Rank Questionnaire (SUPR-Q) results suggested that the interface was perceived as clear and simple, though moderately attractive. Through the Think-Aloud Protocol, experts identified interface elements requiring refinement, mainly related to visual information, navigation,

and feedback. Overall, the findings suggest that DanceMove presents satisfactory usability, while highlighting areas for design improvement.

Keywords: User Interface Design, Foot Interaction, Dance Game.

1 Introduction

Foot interaction has shown considerable potential as an input method in a range of digital solutions. In recent years, several systems have incorporated this form of interaction, including pressure-sensitive floors [1, 2], augmented reality games [3–5], and accessibility tools for users with hand impairments [6]. In these contexts, foot-based interaction has been explored as an alternative input modality capable of expanding interaction possibilities beyond traditional hand-based controls.

When applied to solutions designed for older adults, particularly digital games, foot interaction has been associated with positive outcomes, as this form of interaction encourages physical movement, activates multiple muscle groups, and supports active ageing [7]. Among these solutions, dance games have been widely used in health and care contexts to promote physical and cognitive engagement among older adults [8–10]. By combining music, movement, and visual cues through a dance mat, these games stimulate balance, coordination, attention, and cognitive processing. However, most commercially available dance games are designed for younger audiences and often feature fast-paced rhythms and visually complex interfaces [11], which may not adequately accommodate the motor, cognitive, literacy, and technology-related challenges commonly experienced by older adults [12].

To address these limitations, DanceMove was developed within the "Smart & Healthy Ageing through People Engaging in Supportive Systems" (SHAPES, Horizon 2020 Grant Agreement-857159) by the University of Aveiro. DanceMove is a dance-based digital game that promotes healthy ageing through foot-based interaction and an accessible, socially engaging user interface tailored to older adults. Previous studies [13, 14] have demonstrated that DanceMove was perceived by older adults as a pleasant and motivating digital solution, with high levels of usability, acceptance, and user satisfaction, in addition to being feasible and engaging for use in home environments. These results suggest that DanceMove has great potential to promote physical activity, social interaction, and physical and cognitive training among older adults, as well as helping to overcome common barriers associated with the adoption of digital technologies in this age group.

Despite these results, further research is still needed to better understand how the user interface of dance games with foot interaction can effectively meet the interaction needs of older adults. In particular, it remains important to assess whether interface elements, such as visual layout, navigation structure, colour schemes, and interaction feedback, are suitable for the target users. To fill this research gap, the present study reports and discusses the results of an expert evaluation of the DanceMove user interface. The evaluation was conducted with experts in design and communication sciences and aimed to assess the usability and visual suitability of the interface, considering the characteristics and requirements of the senior target audience.

2 Methods

Previous research identified 244 user interface design guidelines [1-3] to support the development of digital solutions for older adults within the SHAPES project. These guidelines guided the design of the DanceMove interface, ensuring a more inclusive and intuitive experience aligned with the motor, cognitive, and sensory characteristics common among this user group. Earlier work [13] details the interface design process, highlighting the main guidelines followed (e.g., recommendations for multiple interaction paradigms, guidance specifically for foot-based interaction, and principles directed at older adults) as well as the core functional requirements of the dance game, including key screens, navigation structure, and expected features.

The developed interface was evaluated by a group of experts in design and communication sciences to assess usability and determine whether the aesthetic qualities of the game aligned with the needs and preferences of older adults. Most participants had taken part in previous evaluations [16], with a few additional specialists included due to scheduling constraints. Individual test sessions with a playable DanceMove prototype were conducted in person at the University of Aveiro between February and May 2022.

During each session, experts first explored the interface freely and then played one song at each difficulty level (easy, medium, and hard). The Think-Aloud Protocol was used throughout the evaluation to collect qualitative data on navigation and interaction issues in real time.

At the end of the session, each participant completed a 12-item questionnaire combining elements of the SUS and the SUPR-Q. The SUS [18] is a widely used tool for evaluating perceived usability, comprising ten items that cover Nielsen's five usability attributes [19]: learnability, efficiency, memorability, errors, and satisfaction. Participants rated the items on a five-point scale, generating an overall usability score ranging from 0 to 100. The SUPR-Q [20] similarly uses a five-point agreement scale, and includes eight standardised items assessing usability, trust, loyalty, and appearance. Because both instruments contain overlapping usability constructs, items were examined for redundancy. As a result, only two SUPR-Q items — Item 7 (“I find the website attractive”) and Item 8 (“The website has a clean and simple presentation”) — were retained, as they assess visual aspects not covered by the SUS; the remaining items were excluded.

3 Results and Discussion

The expert tests were conducted with six women and four men, with a mean age of 40.7 years ($SD = 8.32$). All participants were researchers in the field of user interface design ($n = 10$, 100%), with academic backgrounds in design ($n = 6$, 60%) and communication sciences and technologies ($n = 4$, 40%). Individual sessions lasted an average of 24 minutes ($SD = 4.53$), during which participants freely navigated the interface, played one song from each of the three difficulty levels, and completed the 12-item

questionnaire. Participant (P) responses to the SUPR-Q items (I) are presented in Table 1.

Table 1. Results of the questionnaire items of the SUPR-Q method.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
I1	3	3	4	4	4	3	3	4	2	3
I2	4	3	5	3	5	4	4	4	4	4

Because calculating the final SUPR-Q score requires the use of all eight items, and only the two appearance-related items were included in this study, the results were analysed individually rather than aggregated into a single score. For the first item, most experts selected the neutral option (score 3), indicating that they found the interface neither particularly attractive nor unattractive ($n = 5$, 50%). For the second item, the majority agreed (score 4) or strongly agreed (score 5) that the interface has a clean and simple presentation ($n = 8$, 80%).

Regarding the SUS items, the participants' raw scores were converted following Brooke's [18] guidelines. For odd-numbered items (1, 3, 5, 7 and 9), the score contribution corresponds to the position on the scale minus one, whereas for even-numbered items (2, 4, 6, 8 and 10), the contribution is calculated as five minus the position on the scale. The adjusted scores are then summed and multiplied by 2.5 to produce the final SUS score. As shown in Table 2, among the ten experts, SUS results ranged from low-acceptable to excellent, with the lowest score recorded at 52.5 and the highest at 97.5. The final average questionnaire score is 75 ($SD = 14.2$), which places it within the range typically interpreted as acceptable to good usability for expert users, given that scores below 60 are generally considered unacceptable and those above 70 acceptable.

Table 2. Results of the questionnaire items of the SUS method.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
I1	2	1	4	2	3	2	2	2	2	3
I2	1	3	4	3	3	4	3	3	3	2
I3	1	3	4	3	3	3	3	3	4	4
I4	3	3	4	3	3	4	4	3	4	3
I5	2	3	4	2	3	3	3	2	3	3
I6	2	3	4	2	1	3	3	2	3	3
I7	3	3	3	1	3	4	2	3	3	4
I8	2	3	4	2	3	4	3	2	3	3
I9	2	3	4	1	2	4	1	3	2	4
I10	3	3	4	2	3	4	4	3	4	4

SUS										
Raw	21	28	39	21	27	35	28	26	31	33
Score										
SUS										
Final	52,5	70	97,5	52,5	67,5	87,5	70	65	77,5	82,5
Score										

Regarding the positive SUS statements, most experts agreed that the game is easy to play ($n = 9$, 90%), that its functionalities are well integrated ($n = 7$, 70%), and that users would learn to play DanceMove quickly ($n = 8$, 80%). Fewer participants expressed interest in frequent use ($n = 3$, 30%) or reported feeling fully confident when using the solution ($n = 5$, 50%). Concerning the negative SUS statements, most experts disagreed that the game is unnecessarily complex ($n = 8$, 80%), inconsistent ($n = 6$, 60%), very difficult to play ($n = 7$, 70%), or that it would require external help to use ($n = 10$, 100%) or extensive prior learning ($n = 9$, 90%).

Qualitative data from the Think-Aloud Protocol complemented the SUS and SUPR-Q results by revealing experts' perceptions of DanceMove's usability and interface design. Several comments addressed visual clarity and information density, helping to explain the neutral ratings observed for the SUPR-Q appearance item. Experts highlighted excessive or repetitive content in the dance mat tutorial screen, difficulty selection screen, and game tutorials screen, recommending simplification and clearer organisation.

Additional suggestions focused on improving interface clarity and consistency, including reorganising menu options, standardising iconography, removing unnecessary visual elements, and enriching the music selection screen with track duration and audio previews. During gameplay, experts proposed displaying the score in real time, refining the countdown bar, and improving score-related screens with clearer visual cues. These recommendations emphasise the need to reduce cognitive load while preserving informative feedback, especially for older adult users.

Recurring usability issues were identified as potentially influencing SUS scores: the help bar, the game screen, difficulty levels and arrow speed, and in-game feedback. While the help bar was considered useful, its inconsistent placement across screens was seen as disruptive, highlighting the importance of spatial consistency. The game screen was also criticised for lacking clear instructions for first-time users, leading experts to suggest a short tutorial or informational overlay, along with a longer countdown supported by audio cues.

Finally, experts noted that the medium and hard difficulty levels were excessively fast and contained too many arrows, potentially limiting accessibility for older adults, and stressed the importance of aligning music tempo with difficulty level. In-game feedback messages were generally viewed as motivating but also as a possible source of visual overload, with audio feedback suggested as a more suitable alternative. Overall, these qualitative findings reinforce the quantitative results, indicating good usability alongside clear opportunities for design refinement.

4 Conclusion

This study evaluated the usability and accessibility of the DanceMove user interface through an expert-based assessment combining quantitative and qualitative methods. Overall, the results indicate satisfactory usability, with experts recognising the game's potential to support active and healthy ageing through foot-based interaction. However, the findings also highlight the need for design refinements, particularly regarding information density, interaction pacing, and consistency across interface elements. Addressing these issues may further enhance accessibility, user confidence, and overall engagement for older adults.

It is important to acknowledge that the number of experts involved in the evaluation was relatively small ($n = 10$), which may limit the results. However, expert evaluations are commonly used in the initial phases of interface evaluation and can provide valuable information for identifying usability problems and guiding design improvements. Therefore, future investigations should involve larger-scale studies and the application of additional evaluation techniques, such as methods focused on exploring the user experience when interacting with the game. Adding a larger number of experts can contribute to a more comprehensive assessment of the interface design and usability aspects, while studies involving older adults themselves can provide important information on how target users experience DanceMove.

By combining expert-based usability evaluations with user-centered experience evaluations in future research, it will be possible to further refine the design of DanceMove and strengthen its potential as an accessible and engaging digital solution to promote active and healthy aging.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Schmidt, D., Frohnhofen, J., Knebel, S., Meinel, F., Perchyk, M., Risch, J., Striebel, J., Wachtel, J., Baudisch, P.: Ergonomic Interaction for Touch Floors. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 3879–3888. ACM, New York, NY, USA (2015). <https://doi.org/10.1145/2702123.2702254>.
2. Sangsuriyachot, N., Sugimoto, M.: Novel interaction techniques based on a combination of hand and foot gestures in tabletop environments. In: Proceedings of the 10th asia pacific conference on Computer human interaction. pp. 21–28. ACM, New York, NY, USA (2012). <https://doi.org/10.1145/2350046.2350053>.
3. Lv, Z., Halawani, A., Feng, S., ur Réhman, S., Li, H.: Touch-less interactive augmented reality game on vision-based wearable device. *Pers. Ubiquitous Comput.* 19, 551–567 (2015). <https://doi.org/10.1007/s00779-015-0844-1>.
4. Lv, Z., Halawani, A., Feng, S., Li, H., Réhman, S.U.: Multimodal Hand and Foot Gesture Interaction for Handheld Devices. *ACM Transactions on Multimedia*

- Computing, Communications, and Applications. 11, 1–19 (2014). <https://doi.org/10.1145/2645860>.
5. Lv, Z., Lal Khan, M.S., Ur Réhman, S.: Hand and foot gesture interaction for handheld devices. In: Proceedings of the 21st ACM international conference on Multimedia. pp. 621–624. ACM, New York, NY, USA (2013). <https://doi.org/10.1145/2502081.2502163>.
 6. Carrozza, M.C., Persichetti, A., Laschi, C., Vecchi, F., Lazzarini, R., Vacalebri, P., Dario, P.: A Wearable Biomechatronic Interface for Controlling Robots with Voluntary Foot Movements. IEEE/ASME Transactions on Mechatronics. 12, 1–11 (2007). <https://doi.org/10.1109/TMECH.2006.886250>.
 7. Warburton, D.E.R., Bredin, S.S.D., Horita, L.T.L., Zbogar, D., Scott, J.M., Esch, B.T.A., Rhodes, R.E.: The health benefits of interactive video game exercise. Applied Physiology, Nutrition, and Metabolism. 32, 655–663 (2007). <https://doi.org/10.1139/H07-038>.
 8. Laver, K., Ratcliffe, J., George, S., Burgess, L., Crotty, M.: Is the Nintendo Wii Fit really acceptable to older people?: a discrete choice experiment. BMC Geriatr. 11, 64 (2011). <https://doi.org/10.1186/1471-2318-11-64>.
 9. Lange, B., Requejo, P., Flynn, S., Rizzo, A., Valero-Cuevas, F., Baker, L., Winstein, C.: The Potential of Virtual Reality and Gaming to Assist Successful Aging with Disability. Phys. Med. Rehabil. Clin. N. Am. 21, 339–356 (2010). <https://doi.org/10.1016/j.pmr.2009.12.007>.
 10. Hanneton S., V.A., Hanneton, S.: Coaching the Wii: Evaluation of a physical training experiment assisted by a video game. In: 2009 IEEE International Workshop on Haptic Audio visual Environments and Games. pp. 54–57. IEEE (2009). <https://doi.org/10.1109/HAVE.2009.5356134>.
 11. Smith, S.T., Sherrington, C., Studenski, S., Schoene, D., Lord, S.R.: A novel Dance Dance Revolution (DDR) system for in-home training of stepping ability: basic parameters of system use by older adults. Br. J. Sports Med. 45, 441–445 (2011). <https://doi.org/10.1136/bjsm.2009.066845>.
 12. Patsoule, E., Koutsabasis, P.: Redesigning websites for older adults: a case study. Behaviour & Information Technology. 33, 561–573 (2014). <https://doi.org/10.1080/0144929X.2013.810777>.
 13. Diehl, C., Silva, T., Ferraz, J., Almeida, A., Martins, A., Ribeiro, Ó., Santinha, G., Rocha, N., Silva, A.: DanceMove: Promoting Social Interaction and Healthy Ageing Through Foot-Interaction. In: Proceedings of the 11th International Conference on Information and Communication Technologies for Ageing Well and e-Health. pp. 421–432. SCITEPRESS - Science and Technology Publications (2025). <https://doi.org/10.5220/0013504100003938>.
 14. Martins, A.I., Diehl, C., Silva, T., Almeida, A.M., Santinha, G., Ribeiro, Ó., Rocha, N.P., Silva, A.G.: From ideation to real context of use of a digital solution to promote physical and cognitive training for older adults. Sci. Rep. 15, 1867 (2025). <https://doi.org/10.1038/s41598-024-83534-w>.
 15. Diehl, C.: Foot Interaction: Design Guidelines for Older Adults, <http://hdl.handle.net/10773/42454>, (2024).

16. Diehl, C., Martins, A., Almeida, A., Silva, T., Ribeiro, Ó., Santinha, G., Rocha, N., Silva, A.G.: Defining Recommendations to Guide User Interface Design: Multimethod Approach. *JMIR Hum. Factors*. 9, e37894 (2022). <https://doi.org/10.2196/37894>.
17. Silva, A., Rocha, N., Almeida, A.M., Silva, T., Ribeiro, Ó., Santinha, G., Tavares, R.: Deliverable 5.1 – SHAPES User Experience Design and Guidelines and Evaluation. (2019).
18. Brooke, J.: SUS: A quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., and McClelland (eds.) *Usability evaluation in industry*. pp. 189–194. Taylor & Francis, London (1996).
19. Nielsen, J.: 10 Usability Heuristics for User Interface Design, <https://www.nngroup.com/articles/ten-usabilityheuristics/>, last accessed 2021/12/20.
20. Sauro, J.: SUPR-Q: A Comprehensive Measure of the Quality of the Website User Experience. *J. Usability Studies*. 10, 68–86 (2015).

Digital Accessibility



Revisión Sistemática sobre Entornos Virtuales 3D en la Educación de Personas con Discapacidad Auditiva

Adriana Silvia Fachal ^[0000-0002-6379-2856] and María José Abásolo ^{1,2[0000-0003-4441-3264]} and Cecilia Verónica Sanz ^{1[0000-0002-9471-0008]}

¹ Instituto Investigación en Informática LIDI, Facultad de Informática – Universidad Nacional de La Plata, La Plata, Argentina

² Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, CICPBA
afachal@hotmail.com, {mjabasolo, csanz}@lidi.info.unlp.edu.ar

Resumen. Se llevó a cabo una revisión sistemática con el objetivo de identificar, describir y analizar experiencias educativas mediadas por tecnología 3D dirigidas a personas con discapacidad auditiva. Se seleccionaron 9 estudios relevantes. Los resultados evidencian el uso de entornos virtuales tridimensionales como herramientas accesibles e inmersivas que favorecen la comunicación visual, la autonomía, la motivación y la participación activa de los estudiantes sordos. Asimismo, se identificaron barreras técnicas relacionadas con dispositivos y condiciones de uso. Las experiencias reportadas implementan diversas estrategias pedagógicas y herramientas digitales, y destacan el potencial de los entornos virtuales tridimensionales para fortalecer procesos educativos inclusivos. Esta revisión permite visibilizar tendencias, vacíos y desafíos para futuras investigaciones en el campo de la educación accesible apoyada por tecnologías inmersivas.

Keywords: Discapacidad Auditiva, Educación Inclusiva, Entornos Virtuales 3D, Revisión Sistemática, Tecnologías Inmersivas.

1 Introducción

En los últimos años, los entornos virtuales tridimensionales (EV3D) han adquirido relevancia como herramientas pedagógicas que permiten potenciar experiencias educativas innovadoras, interactivas e inmersivas. Estas tecnologías, que recrean espacios y situaciones a través de escenarios simulados en tres dimensiones, han demostrado un notable potencial en el desarrollo de propuestas accesibles para personas con diversas condiciones, entre ellas, la discapacidad auditiva.

La educación inclusiva, concebida como un derecho y una práctica que reconoce y valora la diversidad, requiere recursos tecnológicos que respondan a las particularidades de cada grupo de estudiantes.

A pesar del creciente interés en el desarrollo de propuestas educativas con tecnologías inmersivas, aún persiste una dispersión en la sistematización del conocimiento existente sobre experiencias dirigidas a personas sordas o con discapacidad auditiva. La literatura evidencia trabajos aislados, con enfoques, objetivos y herramientas va-

riadas, que dificultan la consolidación de buenas prácticas o líneas comunes de acción

En el marco de una investigación doctoral [1] [2], se llevó a cabo una revisión sistemática con el fin de identificar, describir y analizar experiencias educativas que integren EV3D en propuestas inclusivas para personas con discapacidad auditiva.

2 Metodología

2.1 Protocolo de Revisión

La Revisión Sistemática (RS) desarrollada sigue los lineamientos sugeridos por Kitchenham (2004).

Estas etapas fueron adaptadas al campo de estudio de los EV3D aplicados a la educación de personas con discapacidad auditiva.

En este primer paso se especifican preguntas de investigación que deben ser respondidas.

1- ¿Qué experiencias se han desarrollado en EV3D para personas con discapacidad auditiva y cuáles han sido sus resultados?

2- ¿Qué posibilidades y barreras ofrecen los EV3D para la comunicación y educación de las personas con discapacidad auditiva?

3- ¿Qué estrategias, herramientas y metodologías se han empleado para mediar procesos educativos en EV3D con personas con discapacidad auditiva.

2.2 Estrategia de Búsqueda

Se realizaron búsquedas en bases de datos científicas como Google Académico, IEEE Xplore Digital Library y ACM Digital Library, acotadas a publicaciones relacionadas con experiencias educativas con estudiantes con discapacidad auditiva en EV3D y con un rango temporal comprendido entre 2003 hasta 2024.

Se seleccionaron trabajos publicados en español e inglés. Las palabras clave utilizadas surgieron a partir de puntos claves relacionados a esta RS tales como sordos (discapacidad/discapacitado auditiva/o), plataformas (Open Simulator, Second Life), entorno virtual 3D (realidad virtual, metaverso, mundo virtual, Web3D, avatar, tridimensional, 3D).

Se ha acotado la búsqueda con el término de realidad virtual pero ante la amplitud y dispersión de los resultados obtenidos, se procedió a refinar la estrategia combinando los términos clave con las palabras “enseñanza” o “aprendizaje”. De este modo, se logró filtrar los resultados hacia experiencias educativas concretas en EV3D dirigidas a personas con discapacidad auditiva.

2.3 Filtrado de Artículos

Se incluyeron:

- Artículos en revistas, capítulos de libros derivados de actas de congresos, tesis, capítulos de tesis o reportes de investigación que describan experiencias realizadas con personas con discapacidad auditiva en entornos virtuales tridimensionales

Se excluyeron:

- CE1. Artículos propios; artículos duplicados o diferentes artículos que relatan la misma experiencia, en cuyo caso se selecciona el más reciente; o aquellos para los que no se consigue el documento completo;
- CE2. Artículos de revisión o documentos que no corresponden a artículos de investigación (por ejemplo libros completos o actas completas de congresos)
- CE3. Artículos que no se desarrollan en entorno virtuales 3D: tratan de experiencias en ambientes 2D; contienen sólo animación 3D en lugar de un ambiente 3D (por ejemplo de un avatar); de realidad aumentada; reconocimiento de gestos
- CE4. Artículos que se enfocan en experiencias no visuales (por ej. hápticas)
- CE5. Artículos que no relatan experiencias en ambientes virtuales tridimensionales con personas con discapacidad auditiva

La Tabla 1 muestra la cantidad de artículos filtrados y resume los resultados obtenidos en cada base de datos, incluyendo la cantidad de artículos recuperados, descartados y finalmente seleccionados.

	Cant	CE1	CE2	CE3	CE4	CE5	Eliminados	Seleccionados
Google Scholar	23	8	3	5	1	1	18	5
ACM	22	3	1	16	0	1	21	1
IEEE	31	3	4	16	0	5	28	3
Total	70	13	7	33	1	7	61	9

Tabla 1. Resultado de Cantidad de Artículos obtenidos para las cadenas de búsquedas.

3 Resultados

La revisión sistemática permitió identificar nueve estudios relevantes publicados entre 2003 y 2023. La Tabla X presenta una síntesis de los trabajos seleccionados.

Ref	Autores (año)	Título
[3]	Passig & Eden (2003)	Cognitive intervention through virtual environments among deaf and hard-of-hearing children
[4]	Adamo-Villani & Wright (2007)	SMILE: an immersive learning game for deaf and hearing children
[5]	Adamo-Villani & Wilbur (2008)	Effects of platform (immersive versus non-immersive) on usability and enjoyment of a virtual learning environment for deaf and hearing children
[6]	Lin, Wang, Hung & Lin (2010)	Reducing cognitive load through virtual environments among hearing-impaired students
[7]	Ferreira Brega, Rodello, Colombo Dias & Guimarães (2014)	A virtual reality environment to support chat rooms for hearing impaired and to teach Brazilian Sign Language (LIBRAS)
[8]	Insook (2017)	Prototyping training program in immersive virtual learning environment with head mounted displays and touch-less interfaces for hearing-impaired learners
[9]	Izaguirre, Abásolo & Collazos (2020)	An experiential learning approach to learning manual communication through a virtual reality environment
[10]	Rahim (2022)	A virtual reality approach to support Malaysian sign language interactive learning for deaf-mute children
[11]	Mystakidis, Theologi-Gouti & Iliopoulos (2023)	STEAM Project Exhibition in the Metaverse for Deaf High School Students' Affective Empowerment

Tabla 2. Estudios seleccionados en la revisión sistemática

3.1 Resultados vinculados a la pregunta P1

¿Qué experiencias se han desarrollado en EV3D para personas con discapacidad auditiva y cuáles han sido sus resultados?

Las experiencias revisadas sobre el reconocimiento de señas y gestos se enfocan en el uso de entornos virtuales 3D con dispositivos específicos, buscando aprovechar prácticas como la rotación espacial y la comunicación mediante avatares señantes. Estas propuestas han sido diseñadas con fines educativos diversos: desde el aprendizaje de lengua de señas, hasta el desarrollo de habilidades de navegación en entornos virtuales, la adquisición de vocabulario y la realización de tareas colaborativas. En todos los

casos, los EV3D permitieron a los estudiantes interactuar con escenarios inmersivos y manipular objetos virtuales en función de objetivos pedagógicos.

Los resultados reportados muestran que, en ciertas tareas, los alumnos con discapacidad auditiva obtuvieron puntuaciones similares o levemente inferiores a las de sus pares oyentes. Sin embargo, se destaca su autonomía y creatividad en la construcción y manipulación de objetos en el entorno. La creación de un metaverso propio generó confianza, motivación para una participación activa y un efecto social positivo en los estudiantes, fomentando la interacción y el trabajo colaborativo.

3.2 Resultados vinculados a la pregunta P2

¿Qué posibilidades y barreras ofrecen los EV3D para la comunicación y educación de las personas con discapacidad auditiva?

Los hallazgos indican que los EV3D ofrecen múltiples posibilidades en el ámbito educativo para estudiantes con discapacidad auditiva. Entre ellas se incluyen la posibilidad de ofrecer experiencias realistas e inmersivas, acceso a escenarios seguros para el desarrollo de tareas de la vida diaria, manipulación de objetos en un entorno virtual y control de las acciones por parte del propio usuario. Estos entornos permiten que los estudiantes avancen a su propio ritmo, repitan acciones y visualicen o perciban de forma tangible elementos y procesos, lo que incrementa su motivación y compromiso con el aprendizaje.

La potencialidad de los EV3D para la comunicación se evidencia en la inclusión de avatares con uso de lengua de señas en el 80% de las experiencias revisadas. Estos avatares fueron programados para representar diferentes lenguas de señas, como la Lengua de Señas Americana (ASL), la Lengua de Señas Brasileira (LIBRAS), la Lengua de Señas de Nueva Zelanda (NZSL) y la Lengua de Señas de Malasia (MSL). Solo una experiencia incorporó actividades que requerían representación dactilológica para el aprendizaje de conceptos. La manipulación de elementos del entorno, combinada con la comunicación visual de los avatares, se destacó como una herramienta de gran valor para la comprensión y la interacción en contextos educativos.

Las barreras identificadas incluyen la incompatibilidad de algunas plataformas con dispositivos móviles, la necesidad de equipos específicos como gafas de realidad virtual o rastreadores de movimiento, y problemas físicos reportados por algunos participantes, tales como mareos o fatiga visual tras el uso prolongado de dispositivos inmersivos.

3.3 Resultados vinculados a la pregunta P3

¿Qué estrategias, herramientas y metodologías se han empleado para mediar procesos educativos en EV3D con personas con discapacidad auditiva?

En cuanto a las tecnologías y plataformas empleadas, se identificó el uso de entornos 3D inmersivos como Oculus Rift y Biblioteca de gráficos web (WebGL), así como aplicaciones de escritorio, sistemas CAVE, pantallas inmersivas y sensores Kinect. Para el modelado y animación de objetos y personajes se utilizaron herramientas co-

mo SketchUp Pro y Blender, además de kits de desarrollo para optimizar la experiencia inmersiva.

Las estrategias pedagógicas implementadas incluyeron el uso de gamificación para el aprendizaje de ortografía y vocabulario, tareas de construcción de objetos virtuales, actividades de navegación en entornos simulados y la utilización de salas de chat colaborativas. Se evaluaron aspectos como comprensión, desempeño, usabilidad, satisfacción y disfrute, comparando el uso de soportes inmersivos frente a soportes más tradicionales como el PC.

Metodológicamente, las experiencias recurrieron a grupos contrastados (experimental y de control), comparativas entre usuarios sordos y oyentes, y estudios de caso con grupos específicos de personas con discapacidad auditiva. Los instrumentos de evaluación empleados fueron entrevistas semiestructuradas y no estructuradas, encuestas, cuestionarios, pruebas de usabilidad, observación directa y escalas pictóricas. También se utilizaron técnicas de pensamiento en voz alta para registrar las percepciones de los participantes durante las actividades.

4 Discusión

El análisis de los nueve estudios seleccionados revela un creciente interés por el uso de EV3D en la educación de personas con discapacidad auditiva, aunque este campo aún se encuentra en una fase exploratoria. Si bien las experiencias revisadas evidencian un potencial significativo de estas tecnologías para favorecer la comunicación, la motivación y la participación activa, su aplicación presenta desafíos que requieren ser abordados para lograr una implementación más amplia y efectiva.

En relación con las posibilidades, los EV3D se distinguen por su capacidad para ofrecer entornos inmersivos, realistas y controlados que facilitan la manipulación de objetos virtuales y la representación visual de conceptos abstractos. La incorporación de avatares con uso de lengua de señas en la mayoría de los estudios constituye un aporte relevante para la accesibilidad comunicativa, permitiendo que las interacciones se desarrollen en un marco inclusivo y adaptado a las necesidades de los usuarios sordos. Sin embargo, la diversidad de lenguas de señas empleadas y la ausencia de estándares unificados limitan la interoperabilidad y la transferencia de estas soluciones a otros contextos educativos.

El análisis comparativo muestra que gran parte de las experiencias se llevaron a cabo en entornos experimentales o de laboratorio, con grupos reducidos y en periodos cortos de intervención. Esto, si bien permite un control riguroso de las variables, restringe la posibilidad de evaluar el impacto sostenido de los EV3D en el aprendizaje. La falta de estudios longitudinales y de implementaciones en contextos educativos formales constituye una limitación recurrente identificada en la literatura revisada.

Asimismo, se observa una marcada heterogeneidad en las tecnologías y metodologías utilizadas, desde plataformas comerciales hasta entornos desarrollados ad hoc, así como en las estrategias pedagógicas y los criterios de evaluación. Esta diversidad enriquece el panorama de posibilidades, pero también dificulta la comparación de resultados y la generación de recomendaciones sólidas para la adopción de EV3D en la educación inclusiva.

En cuanto a las barreras, además de las limitaciones técnicas vinculadas a compatibilidad y requerimientos de hardware, algunos estudios reportan inconvenientes físicos como mareos o fatiga visual, lo que indica la necesidad de considerar criterios de ergonomía y salud en el diseño e implementación de experiencias inmersivas. Estos aspectos, poco abordados en la literatura, representan un campo de investigación complementario que podría mejorar la experiencia de uso y la aceptación de estas tecnologías.

Finalmente, la revisión pone en evidencia que la mayoría de las experiencias analizadas se centran en la interacción entre personas sordas y el entorno virtual, mientras que son escasas las propuestas que promueven la interacción conjunta entre personas sordas y oyentes en escenarios inclusivos. Explorar estas dinámicas podría aportar conocimientos valiosos sobre el potencial de los EV3D para favorecer el aprendizaje colaborativo, la cooperación y la sensibilización en comunidades educativas diversas.

Del análisis de la literatura revisada se identifican varias líneas de trabajo que podrían orientar investigaciones futuras en el campo de los EV3D aplicados a la educación de personas con discapacidad auditiva. En primer lugar, se evidencia la necesidad de desarrollar estudios longitudinales que permitan evaluar de forma sostenida el impacto de la utilización de EV3D en el aprendizaje, la comunicación y la inclusión social de las personas sordas. La mayoría de las experiencias reportadas corresponden a estudios de corta duración, lo que limita la comprensión de sus efectos a largo plazo.

En segundo lugar, se observa una escasa comparación sistemática entre distintas plataformas y tecnologías inmersivas, lo que dificulta establecer qué herramientas ofrecen mejores condiciones de accesibilidad, usabilidad y eficacia pedagógica para esta población. Investigaciones comparativas podrían contribuir a generar recomendaciones más precisas para la selección de entornos y dispositivos.

Otra línea de interés consiste en la integración de estándares y protocolos de accesibilidad específicos para EV3D, incluyendo guías para la representación de lengua de señas y la adaptación de interfaces visuales. Esto permitiría avanzar hacia la construcción de entornos virtuales inclusivos desde el diseño, evitando barreras técnicas y comunicativas.

Asimismo, resulta pertinente explorar la transferencia de experiencias educativas desarrolladas en EV3D a contextos formales de enseñanza, dado que la mayoría de los estudios se realizaron en entornos experimentales o de carácter informal. Ello podría implicar la incorporación de EV3D en planes de estudio, evaluando su impacto en la adquisición de competencias curriculares.

Finalmente, se advierte un campo poco explorado en torno a la interacción entre personas sordas y oyentes en entornos virtuales inclusivos, que permitiría estudiar cómo los EV3D pueden favorecer la cooperación, la sensibilización y el aprendizaje colaborativo en contextos educativos diversos.

5 Conclusiones

La revisión sistemática realizada permitió identificar y analizar nueve estudios que abordaron experiencias educativas mediadas por entornos virtuales tridimensionales EV3D dirigidas a personas con discapacidad auditiva. El análisis evidenció que estos entornos ofrecen oportunidades significativas para la enseñanza y la comunicación, destacándose la potencialidad inmersiva, la manipulación de objetos virtuales y la incorporación de avatares con uso de lengua de señas como elementos facilitadores de la comprensión y la interacción.

Los resultados indican que los EV3D favorecen la autonomía, la motivación y la participación activa de los estudiantes, además de proporcionar escenarios seguros y controlados para el desarrollo de tareas educativas. Sin embargo, persisten barreras vinculadas a la disponibilidad y compatibilidad tecnológica, así como a la necesidad de hardware específico y posibles efectos físicos derivados del uso prolongado de dispositivos inmersivos.

Se constató una diversidad de tecnologías, estrategias pedagógicas y metodologías de evaluación, lo que refleja el carácter experimental y exploratorio de gran parte de las experiencias revisadas. Esta heterogeneidad, si bien enriquece el campo, dificulta la comparación de resultados y la definición de estándares de accesibilidad y usabilidad.

En conjunto, los hallazgos ponen de manifiesto que los EV3D constituyen una herramienta valiosa para la educación inclusiva de personas con discapacidad auditiva, pero requieren mayor consolidación metodológica, estudios longitudinales y validaciones en contextos educativos formales para alcanzar su máximo potencial.

Referencias

1. Fachal, A.S (2024). Posibilidades pedagógicas de los entornos virtuales 3D en el acompañamiento del aprendizaje de personas con discapacidad auditiva. Universidad Nacional de La Plata. <http://sedici.unlp.edu.ar/handle/10915/174350>. <https://doi.org/10.35537/10915/174350>
2. Fachal, A., Abásolo, M. J., Sanz, C. V. (2024). Experiencia educativa con personas con discapacidad auditiva mediante el uso de entornos virtuales 3D en OpenSim. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología (TE&ET)*, (38), e14-e14. ISSN 1850-9959
3. Passig, D., & Eden, S. (2003). Cognitive intervention through virtual environments among deaf and hard-of-hearing children. *European Journal of Special Needs Education*, 18(2), 173-182. <https://doi.org/10.1080/0885625032000078961>
4. Adamo-Villani, N., & Wright, K. (2007). SMILE: an immersive learning game for deaf and hearing children. *SIGGRAPH '07: ACM SIGGRAPH 2007 educators program*, (pp 17).
5. Adamo-Villani, N., & Wilbur, R. B. (2008). Effects of Platform (Immersive versus Non-immersive) on Usability and Enjoyment of a Virtual Learning Environment for Deaf and Hearing Children. In *EGVE (Posters)*.

6. Lin, C. Y., Wang, L. C., Hung, P. H., & Lin, C. C. (2010). Reducing cognitive load through virtual environments among hearing-impaired students. In 2010 Second Pacific-Asia Conference on Circuits, Communications and System, (págs. 183-186).
7. Ferreira Brega, J. R., Rodello, I. A., Colombo Dias, D. R., & Guimarães, M. a. (2014). A virtual reality environment to support chat rooms for hearing impaired and to teach Brazilian Sign Language (LIBRAS). 11th International Conference on Computer Systems and Applications (AICCSA), (pp. 433-440). Doha, Qatar.
8. Insook, H. A. (2017). Prototyping training program in immersive virtual learning environment with head mounted displays and touchless interfaces for hearing-impaired learners. Educational Technology International, (pp 49-71).
9. Izaguirre, J., Abásolo, M., Collazos, C. A.: An Experiential Learning Approach to Learning Manual Communication Through a Virtual Reality Environment. In: Applications and Usability of Interactive Media. Communications in Computer and Information Science (CCIS), vol. 1227, pp. 33–45. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65218-0_3
10. Rahim, N. R. (2022). A virtual reality approach to support malaysian sign language interactive learning for deaf-mute children. Journal of pharmaceutical negative results, (págs. 4962-4974).
11. Mystakidis, S., Theologi-Gouti, P., & Iliopoulos, I. (2023). STEAM Project Exhibition in the Metaverse for Deaf High School Students' Affective Empowerment: The Power of Student Museum Exhibitions in Social Virtual Reality. In International Conference on Immersive Learning (pp. 239-249). Cham: Springer Nature Switzerland.

ISDB-T One-Seg Receiver with Automatic Integration of Ecuadorian Sign Language (LSEC) Interpretation Using Neural Networks

Gonzalo Olmedo, Bryan Carrillo, Víctor Zurita, and Nancy Paredes

Universidad de las Fuerzas Armadas - ESPE
gfolmedo@espe.edu.ec

Abstract. This work presents the development of a software-defined radio (SDR) system capable of automatically integrating Ecuadorian Sign Language (LSEC) into Digital Terrestrial Television (DTT) content. The system employs an ADALM-Pluto device and GNU Radio for signal reception, demodulation, and extraction of the transport stream (TS). An external Python module was implemented to perform audio-to-text transcription using the Faster-Whisper model and to detect matches with a sign-language dictionary composed of 3D-avatar video clips. The prototype overlays the corresponding gestural sequences onto the received video stream in real time. Comparative tests with different transcription models yielded latencies ranging from 12 s to 48 s per segment, depending on model complexity and hardware performance. The results demonstrate the feasibility of developing low-cost accessibility solutions for individuals with hearing disabilities in Ecuador.

Keywords: Digital Terrestrial Television (DTT) · ISDB-T One-Seg · Software-Defined Radio (SDR) · Ecuadorian Sign Language (LSEC) · Faster-Whisper · Accessibility Technologies · 3D Avatar Integration

1 Introduction

Digital Terrestrial Television (DTT) plays a central role in mass communication and public information services. In Ecuador, the adoption of the Japanese Brazilian ISDB-Tb standard marked the transition from analog to digital broadcasting, supported by technical studies and international cooperation [1] [2]. This process was later consolidated through national deployment plans focused on infrastructure development and service expansion [3].

Despite these advances, accessibility services for deaf audiences remain limited. Although national regulations promote the inclusion of Ecuadorian Sign Language (LSEC) in television programming, its implementation is still largely restricted, leaving significant gaps in access to audiovisual content [4] [5]. Previous research in Ecuador addressed this issue by embedding prerecorded LSEC content and, more recently, by introducing 3D avatar based sign language representations [6] [7]. Similar international initiatives, such as VLibras in Brazil, demonstrated the scalability of avatar based interpretation [8]. However, these

proaches are typically applied during content production rather than at the receiver.

At the same time, software defined radio (SDR) technologies combined with tools like GNU Radio have enabled low cost reception and analysis of ISDB-T signals using devices such as ADALM Pluto [9] [10]. In parallel, advances in automatic speech recognition have produced neural models such as Whisper and Faster Whisper, capable of robust transcription in noisy broadcast environments [11]. Nevertheless, these technologies have not yet been integrated into end user receivers to provide real time sign language interpretation.

This work addresses this gap by presenting an SDR based ISDB-T One Seg receiver that integrates neural audio transcription with dictionary based LSEC avatar rendering at the receiver side. Experimental evaluations assessed transcription accuracy, latency, and system synchronization under live broadcast conditions, demonstrating the feasibility of combining SDR technology and neural ASR models to enhance accessibility for deaf and hard of hearing audiences in Ecuador.

2 Methodology

Software defined radio (SDR) platforms such as RTL-SDR and ADALM Pluto enable low cost reception and analysis of digital television signals using open source tools like GNU Radio. Prior studies have demonstrated that these devices can reliably demodulate ISDB-T One Seg transmissions and extract transport streams for experimental applications [9] [10].

In parallel, accessibility research has advanced through the use of virtual avatars for sign language interpretation. Systems such as VLibras show that gesture rendering through 3D avatars is a viable approach for supporting deaf audiences [8]. Ecuadorian initiatives have similarly developed LSEC video integration and expanded avatar gesture libraries for multimedia content [6] [7]. However, these efforts operate at the content production stage rather than within the receiver.

Recent progress in transformer based automatic speech recognition (ASR) models, including Whisper and Faster Whisper, provides accurate multilingual transcription suitable for real time applications [11]. The combination of SDR reception, neural transcription, and avatar based gesture rendering has not yet been implemented for One Seg receivers.

2.1 System Overview

The proposed system integrates SDR based signal acquisition, ASR transcription, LSEC dictionary matching, and avatar rendering into a single receiver side pipeline.

Signal Acquisition and Demodulation An ADALM Pluto SDR captures the ISDB-T One Seg signal. A GNU Radio flowgraph performs synchronization,

OFDM demodulation, deinterleaving, and error correction to extract a continuous transport stream.

Transport Stream Processing A Python module segments the transport stream into short intervals, decodes each segment, and extracts the audio component. This design maintains a stable processing load and preserves temporal alignment with the video.

Neural Audio Transcription The audio segments extracted from the transport stream are transcribed using Faster-Whisper. Whisper is a family of open-source automatic speech recognition (ASR) models based on a Transformer encoder-decoder architecture that converts audio representations in log-Mel spectrograms into text. These models were trained on approximately 680,000 hours of audio-text pairs collected from the web, giving them robustness across different languages, accents, and acoustic conditions. Tables 1 and 2 present a comparison of several transcription tools, including cloud-based and locally run solutions. Since the goal of the proposed system is to implement a low-cost solution for LSEC interpretation, tools with no licensing costs were prioritized. In this context, Faster-Whisper was selected for its good transcription quality and, as an optimized implementation of Whisper that uses the CTranslate2 inference engine, it accelerates transcription and reduces memory usage. This allows for efficient execution even on computers with limited computing resources, such as systems without a dedicated graphics card.

Table 1. Comparison of speech transcription tools and execution type

Tool	Execution Type		Cost
Whisper	Local		Open-source
Whisper API	Cloud	Paid	(trial credits available)
Faster-Whisper	Local		Open-source
Vosk	Local		Open-source
Google Speech-to-Text	Cloud	Paid	(trial credits available)

Table 2. Main limitations of speech transcription platforms

Platform	Main Limitation
Whisper	High computational resources required
Whisper API	Requires stable internet connection
Faster-Whisper	Higher processing time on low-resource systems
Vosk	Lower transcription accuracy
Google Speech-to-Text	Internet connection and usage cost

Avatar Gesture Integration The text obtained from each transcribed audio segment is processed using a Python-based algorithm that compares the transcription with labels associated with video clips of a 3D avatar that interprets Ecuadorian Sign Language. Each clip represents a predefined word or short expression. In the current implementation, approximately 150 elements are used, selected as a representative subset to evaluate the system. When a match is detected, the corresponding avatar video clip is overlaid on the original video using FFmpeg, placing it in a fixed region of the frame to avoid interfering with the main broadcast content.

Evaluation Procedure The system was evaluated with live One Seg broadcasts. Metrics included reception stability, transcription accuracy, processing latency, and synchronization across audio, video, and avatar tracks. Additional tests with AI generated audio were used for controlled comparisons.

This streamlined methodology outlines the end to end workflow of the proposed receiver and establishes the foundation for integrating LSEC interpretation directly into digital television content.

3 Results

The system was evaluated using live ISDB-T One Seg broadcasts and controlled audio samples. Performance metrics included reception stability, transcription accuracy, latency per processing stage, and overall synchronization among video, audio, and gesture rendering. Figure 1 illustrates the One Seg receiver developed using software defined radio and implemented in GNU Radio, highlighting the reception and signal processing architecture. Figure 2 shows an example of the reception of the Channel 7 Ecuador TV signal, where automatic speech recognition and sign language interpretation through a virtual avatar are integrated.

3.1 Reception Stability

Tests were conducted at different indoor and outdoor locations using a passive UHF antenna. The ADALM Pluto and GNU Radio flowgraph successfully locked onto the One Seg segment in all test scenarios. Minor pixelation occurred under weak signal conditions, but the audio stream remained decodable for transcription. This behavior is consistent with the robustness characteristics of the ISDB-T standard, which prioritizes audio stability in mobile and low quality reception environments.

3.2 Transcription Latency and Model Comparison

The Faster Whisper model was evaluated in two configurations: a base model optimized for speed and a large model optimized for accuracy. Table 3 summarizes the average processing time per audio segment.



Fig. 1. One Seg receiver developed using software defined radio for ISDB-T One Seg, implemented in GNU Radio.

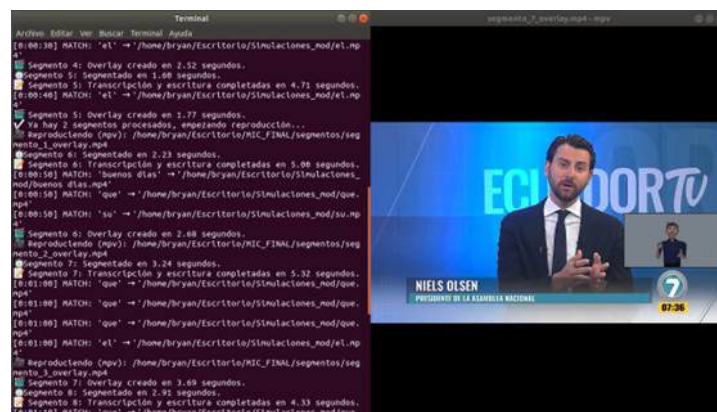


Fig. 2. Reception of the Channel 7 Ecuador TV signal, including automatic speech recognition and sign language interpretation through a virtual avatar.

Table 3. Average transcription latency for different Faster Whisper models

Model	Average Latency (s)	Total Delay (min:s)
Faster Whisper Base	12	00:25
Faster Whisper Large	45 to 48	03:45 to 03:55

The base model delivered significantly lower latency while maintaining acceptable accuracy for short LSEC related phrases. The large model produced more precise transcriptions but introduced delays unsuitable for real time scenarios. These results confirm that model selection must balance accuracy and responsiveness depending on the target application.

3.3 Transcription quality and dictionary matching performance

To evaluate the quality of the subtitles generated by the system, several tests were conducted using different types of television programming. Scenarios with clear audio, particularly in speech, as well as programs containing background noise were considered. Additionally, the evaluation was carried out with the objective of comparing the detection of dictionary elements under different input conditions. Matching performance was evaluated in two scenarios: audio from live broadcast transmissions and speech samples generated using artificial intelligence, whose content mostly includes elements present in the dictionary. Table 4 presents the transcription quality results, while Table 5 shows the obtained detection rates.

Table 4. Transcription quality performance under different audio conditions

Audio condition	Estimated transcription accuracy
Clear speech audio	High quality
Audio with background noise	Medium quality

Table 5. Dictionary matching performance under different audio conditions

Audio Source	Detection Rate	Notes
ISDB-T Live Broadcast	Low	Matches limited to the dictionary content.
AI Generated Audio	High	Good performance in high match detection.

During testing, the system was also evaluated with content containing native Ecuadorian words, such as Kichwa terms and uncommon geographical names. In several cases, the model generated words phonetically similar to the spoken audio, but with spelling errors. On the other hand, the system achieved high detection performance when the transcript contained several entries present in

the dictionary. In live streams, detection rates decreased because the system relies heavily on the audio content and the generated transcript matching the dictionary's contents.

3.4 Gesture Rendering and Synchronization

The gesture rendering engine successfully overlaid avatar animations onto the broadcast video. Synchronization between the interpreted gestures and the transcribed audio remained consistent as long as the delay introduced by the transcription model remained below one minute. For the base model, the avatar appeared within an acceptable temporal window relative to the original speech content. For the large model, the added delay caused noticeable desynchronization.

3.5 Overall System Performance

The complete pipeline was evaluated using continuous One Seg transmissions for periods exceeding thirty minutes. The system operated without interruption and maintained stable video and audio output. Table 6 provides a consolidated view of system performance.

Table 6. Summary of system performance metrics

Metric	Result	Observation
One Seg Reception	Stable	Occasional pixelation under weak signal
TS Extraction	Continuous	No dropped segments
Transcription Latency	12 to 48 s	Depends on model size
Dictionary Matching	Low to High	Requires expanded vocabulary
Avatar Rendering	Successful	Consistent overlay using FFmpeg
End to End Delay	25 s to 4 min	Real time dependent on model selection

4 Discussion and Conclusions

The prototype demonstrates that integrating Ecuadorian Sign Language interpretation directly at the receiver side of an ISDB-T One Seg system is technically feasible using low cost SDR hardware and open source tools. This receiver based approach enables accessibility even when broadcasters do not include sign language services.

Tests showed that transcription latency depends strongly on the selected Faster Whisper model. The base model offers responsiveness suitable for near real time use, while larger models provide higher accuracy at the cost of delays incompatible with live interpretation. This highlights the need to balance accuracy and latency when integrating neural ASR into broadcast workflows.

The results presented show that the dictionary-based matching mechanism performs better when the audio is clear and the transcribed content includes elements present in the dictionary. In contrast, under audio conditions with background noise or pronunciation variability, such as those found in live broadcasts, a decrease in detection rates is observed. These results demonstrate that the system's performance depends on both the quality of the transcription and the coverage of the LSEC dictionary, so expanding the vocabulary and incorporating more flexible linguistic matching methods could improve detection under real-world broadcast conditions.

The system successfully overlaid 3D avatar gestures onto the received video, confirming the potential of avatar based interpretation as a scalable accessibility mechanism. However, improvements in gesture naturalness and expressiveness are needed for broader user acceptance.

Overall, the results confirm that SDR reception, neural transcription, and avatar-based rendering can be integrated into a functional architecture for automatic sign language interpretation in digital television. However, the impact in terms of inclusion and communicational effectiveness has not been empirically validated at this stage of the study. In particular, validation with members of the deaf community was not conducted because there is currently no protocol approved by an ethics committee that authorizes studies involving human participants. As future work, efforts will focus on reducing transcription latency, expanding the LSEC dictionary, improving avatar animations, and conducting evaluations with deaf users to assess the system's communicational performance, following the ethical guidelines established in the Declaration of Helsinki for research involving human subjects.

References

1. CONATEL. Resolución 084-05-CONATEL-2010. Consejo Nacional de Telecomunicaciones, Ecuador (2010).
2. DiBEG. Digital Terrestrial TV Seminar Held in Ecuador. Japan International Cooperation Agency (2015).
3. ARCOTEL. Desarrollo de la Televisión Digital Terrestre en Ecuador. Agencia de Regulación y Control de las Telecomunicaciones (2018).
4. Jiménez, M.: Analysis of Communication Experiences of Deaf People in Ecuadorian Media. *Revista de Comunicación Social*, 15(2), 88–102 (2020).
5. CONADIS. Disability Statistics Report. Consejo Nacional para la Igualdad de Discapacidades, Ecuador (2024).
6. M. Real, M. Santamaría and G. Olmedo, "An architecture for the implementation of the Ecuadorian sign language into digital television system," 2020 IEEE ANDESCON, Quito, Ecuador, 2020, pp. 1-5, doi: 10.1109/ANDESCON50619.2020.9272101.
7. Rivas, R., Maldonado, P., and Córdova, S.: Enhanced 3D Avatar Based Interpretation of Ecuadorian Sign Language for Digital Media. *Revista Ecuatoriana de Telecomunicaciones e Innovación*, 8(2), 55–70 (2024).
8. Silva, R., de Souza, L., Pereira, J., and Almeida, M.: VLibras: An Open Source Accessibility Platform for Brazilian Sign Language. *Journal of Accessibility and Design for All*, 12(1), 45–67 (2022).

9. Larroca, F., and Borrelli, J.: Implementation of an ISDB-T Receiver Using GNU Radio and Low Cost SDR Hardware. *International Journal of Digital Broadcasting Systems*, 14(3), 120–133 (2020).
10. Olmedo, G., Montiel, D., Benavides, N. (2026). Innovative Design and Implementation of a Low-Cost ISDB-T Receiver and Transport Stream Analyzer Using Software-Defined Radio Technology. *Applications and Usability of Interactive TV. jAUTI 2024. Communications in Computer and Information Science*, vol 2514. Springer, Cham. https://doi.org/10.1007/978-3-032-01162-6_2.
11. OpenAI: Whisper and Faster Whisper Models for Multilingual Automatic Speech Recognition. Technical Report, OpenAI Research (2025).



MEO4ALL: Enhancing Interactive Television Accessibility through Voice Interaction and Inclusive Interface Redesign

Rita Oliveira¹[0000-0001-6041-9469], Ana Patrícia Oliveira¹[0000-0003-3234-787X] and Daniel Alves¹

¹ University of Aveiro, Department of Communication and Art, DigiMedia, Aveiro, Portugal
lncs@springer.com

Abstract. Interactive Television (iTV) has evolved into a complex digital ecosystem that integrates linear broadcasting, video-on-demand, electronic program guides, search tools, and personalized recommendations within a single platform. While these developments have expanded media choice, they have also reinforced accessibility barriers for blind and low-vision users who remain largely dependent on visually oriented navigation paradigms. In Portugal, despite gradual improvements in digital television accessibility, visually impaired audiences still face persistent difficulties in navigating iTV services because mainstream interfaces were not originally designed for non-visual interaction.

The MEO4ALL project addresses this problem by integrating a native voice assistant into the MEO Android service through a collaboration between the University of Aveiro and Altice Labs. The solution combines Natural Language Processing (NLP)-based Voice Control via MEO BOTSchool, TalkBack-enabled Voice Over, and Contextual Assistance to enable more intuitive and multimodal interaction. The work builds on previous University of Aveiro initiatives on accessible interactive television and was developed in dialogue with Portuguese associations representing people with visual impairments, ensuring that real user needs informed design decisions.

This paper focuses on the redesign of the MEO4ALL interface, which resulted from a benchmarking analysis of leading commercial voice assistants including Google Assistant, Siri, and Alexa, and from an iterative design process validated through Nielsen's usability heuristics. The new interface adopts a vertical right-side layout, minimalist aesthetics, contextual voice tips, and multimodal feedback including Lottie-based animations and smoother fade transitions. The redesign was prototyped in Figma and implemented in Kotlin within the MEO Android application, enhancing accessibility.

Keywords: Accessibility; Interactive Television (iTV); AI-driven Interaction; User Experience (UX); Inclusive Design

1 Introduction

Interactive Television (iTV) has transformed the way audiences engage with audiovisual content. Contemporary platforms no longer function solely as channels for linear broadcasting; instead, they operate as integrated digital environments that combine live television, streaming libraries, search tools, personalized recommendations, and interactive menus. While this convergence has enriched media consumption for many viewers, it has simultaneously increased interface complexity, often privileging visual interaction over other modalities. For blind and low-vision users, this shift has created new barriers, as navigation within iTV services frequently depends on visual cues, spatial menus, and remote-control interactions that were not originally designed with accessibility in mind [1].

In Portugal, despite gradual improvements in digital television accessibility, significant gaps remain across iTV platforms, particularly regarding navigation, search, and contextual feedback. These limitations do not merely represent technical shortcomings; they also have social implications, as restricted access to iTV can limit participation in cultural life, access to information, and media inclusion [2].

The MEO4ALL project was launched to address these challenges within the MEO Android service¹, one of the most widely used television platforms in Portugal. Developed through a partnership between the University of Aveiro and Altice Labs, MEO4ALL builds on earlier research initiatives at the University of Aveiro dedicated to accessible interactive television and was informed through collaboration with Portuguese associations representing people with visual impairments. These engagements helped identify key needs and expectations of visually impaired users and reinforced the necessity of a native voice assistant tailored to iTV contexts [3].

Rather than focusing on the overall system architecture, this paper concentrates on a central yet underexplored contribution: the redesign of the MEO4ALL user interface. This redesign is treated as a design-research case study that integrates benchmarking, heuristic evaluation, and technical implementation within a real commercial iTV ecosystem.

The paper proceeds by situating the work within research on accessibility and voice interaction in iTV, detailing the MEO4ALL system, explaining the benchmarking and redesign process, and discussing the broader implications of these design decisions for inclusive television design.

2 Background

Research on television accessibility has traditionally focused on content-related features such as audio description, which remain essential for ensuring that audiovisual narratives are perceivable to diverse audiences. At the same time, other accessibility

¹ MEO Android TV service is a digital television service provided by MEO (Altice Portugal), delivered through Android TV devices, allowing users to access live TV channels, on-demand content, and interactive features via an internet connection.

supports embedded in television systems, such as voice interfaces and voice guides, have also been developed to assist visually impaired users in controlling smart TVs and navigating menus [4].

However, making content and interfaces perceivable is only one dimension of accessibility. Equally critical is the usability of the interaction mechanisms themselves. Complex menus, spatial navigation structures, and visually dense overlays can create interaction bottlenecks that prevent users from reaching accessible content efficiently. Recent research in digital audiovisual accessibility indicates that, despite significant technical efforts, substantial gaps persist in how people with visual impairments interact with multimodal systems, including interactive television platforms [1].

Voice interaction has emerged as a particularly promising approach to mitigate some of these usability barriers because it reduces reliance on visual cues and physical remote-control navigation. Commercial voice assistants such as Google Assistant, Siri, and Alexa demonstrate the potential of Natural Language Processing (NLP) to enable task completion through natural language rather than complex visual interactions. A recent systematic review shows that voice user interfaces are rapidly evolving and increasingly recognized to enhance accessibility in multimodal interfaces, yet they still present persistent challenges related to design, context of use, and user diversity [5].

Nevertheless, most commercial voice assistants are primarily optimized for smartphones, smart speakers, or general computing devices, rather than for television ecosystems. iTV involves larger screens, shared viewing contexts, remote controls, and different interaction rhythms, which makes direct adaptation of existing systems insufficient. Recent has also shown how people with visual impairments interact with television environments and document preferences for input modalities such as voice and smartphone control, pointing to the importance of tailored, multimodal solutions for accessible interactive television [6].

Within this context, MEO4ALL positions itself at the intersection of accessibility, UX, and inclusive design by proposing a voice-based solution designed specifically for interactive television rather than adapted from other devices. This distinction is crucial because television remains a shared medium where interaction patterns, spatial layout, and feedback modalities differ significantly from personal mobile devices.

3 The MEO4ALL System Redesign

The MEO4ALL solution integrates three complementary strategies to enhance accessibility within the MEO Android service (Oliveira et al., 2024).

First, Voice Control relies on Natural Language Processing through MEO BOTSchool [7] to interpret user commands such as changing channels, searching for programs, or activating features. This approach prioritizes natural language interaction rather than predefined keywords, aligning with principles of user-centered design that emphasize recognition over recall [8].

Second, Voice Over is implemented via Android's TalkBack screen reader, which provides spoken feedback about on-screen elements and system states [9]. This

integration ensures that blind users can receive auditory descriptions of interface components, reinforcing the multimodal nature of the system.

Third, Contextual Assistance dynamically offers suggestions and guidance based on user actions and system status. For example, if the assistant detects hesitation or repeated unsuccessful commands, it may present spoken tips such as “Try changing channel” or “Activate TalkBack.” This mechanism reduces cognitive load and supports users with low digital literacy or limited familiarity with voice interfaces.

Technically, the solution was implemented in Kotlin to guarantee compatibility with the existing MEO Android application and to facilitate future scalability and maintenance. The architecture thus combines AI-driven interaction with robust software engineering practices suited to commercial deployment.

3.1 Benchmarking Commercial Voice Assistants

A central step in the redesign process was a systematic benchmarking study of three leading commercial voice assistants: Google Assistant [10], Siri [11], and Alexa [12]. The analysis examined layout orientation, visual aesthetics, audio feedback, hardware integration, and accessibility support.

The study revealed that all three assistants predominantly use a horizontal layout positioned at the bottom of the screen. Although visually coherent, this configuration posed a risk of overlapping with existing MEO interface elements such as quick zapping controls, category menus, and search panels. For visually impaired users, this overlap could generate confusion regarding which system was being activated.

Despite this limitation, the benchmarking identified several best practices that informed the MEO4ALL redesign. These included minimalist aesthetics, dark-mode palettes to reduce visual fatigue, clear audio cues signaling system activation, and contextual tips that help users discover available commands. The presence of dedicated remote-control buttons for voice activation was also considered advantageous, as it offers a physical alternative to spoken interaction.

3.2 Mapping Screen Space and Interface Placement

During the ideation phase, the design team mapped five existing MEO interface overlays triggered by remote-control interactions: quick zapping to the left or right (Fig. 1 - I1), category menus activated by the up button (Fig. 1 - I2), detailed channel views triggered by the down button (Fig. 1 - I3), playback controls via the confirmation button (Fig. 1 - I4), and the channel mosaic accessed through the mosaic key (Fig. 1 - I5). This systematic mapping revealed that most overlays occupied horizontal regions at the top or bottom of the screen, while a vertical column on the right side remained consistently unused (Fig. 1).

This observation led to a pivotal design decision: positioning the MEO4ALL assistant vertically on the right side of the screen. The rationale was twofold. First, it avoided visual conflict with existing interface elements, minimizing overlap and cognitive interference. Second, it clearly distinguished the assistant as a separate functional layer rather than an extension of existing menus.

Early sketches explored alternative horizontal layouts across different screen thirds, but these options were ultimately discarded as too intrusive overactive television content. A horizontal search-bar style layout was also considered but abandoned due to its similarity with existing search interfaces. The final decision favored a compact vertical sidebar, balancing visibility, minimal intrusion, and spatial efficiency.

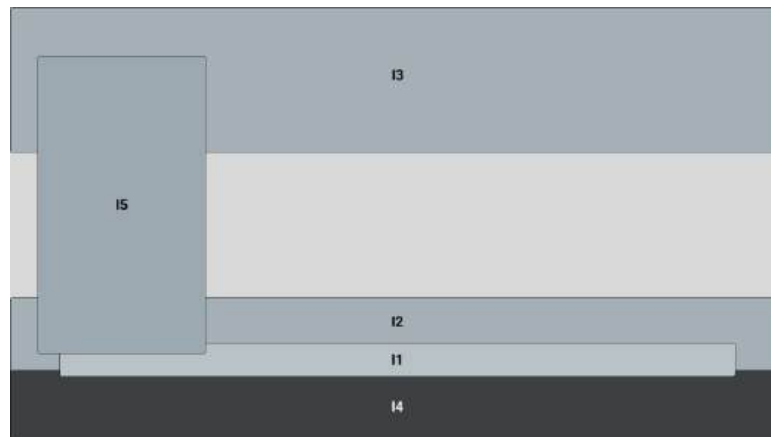


Fig. 1. Mapping of existing MEO overlays and identification of available right-side space.

3.3 Usability Validation through Nielsen's Heuristics

In the absence of immediate large-scale user testing, the team employed Nielsen's [13] ten usability heuristics as a structured evaluation framework for the redesigned interface. This method is widely used in early-stage design to identify potential usability issues before formal empirical testing.

The evaluation highlighted several strengths. The system provided clear feedback through audio cues and visual indicators, such as a microphone animation signaling active listening. The use of natural language supported intuitive interaction, and the visual design remained consistent with other MEO components, reinforcing familiarity.

However, the heuristic analysis also revealed areas for improvement. For instance, the removal of a characteristic audio signal upon activation reduced the visibility of system status, suggesting that this cue should be reinstated. Additionally, error recovery mechanisms were deemed insufficient when speech recognition failed, indicating the need for clearer prompts or alternative interaction paths.

3.4 Implementation

The redesigned interface (Fig. 2) was prototyped in Figma and subsequently implemented in Kotlin within the MEO Android application (Fig. 3). One of the most significant changes was the repositioning of the main layout from a centered position to the right vertical column. This adjustment improved visual organization and clarified functional boundaries.

The “MEO4ALL” label, previously a small caption, was encapsulated within a rounded gray frame to enhance distinction from surrounding elements. Recognized speech text was center aligned to improve readability, while a new tips view was added to present contextual suggestions in a lightweight and non-intrusive manner.

A looping Lottie animation of a microphone was integrated to visually communicate when the system was capturing audio [14]. This multimodal feedback reinforced system transparency and user awareness. Additionally, fragment transitions were changed from sliding animations to fade-in and fade-out effects, reducing visual disruption and creating a smoother experience.

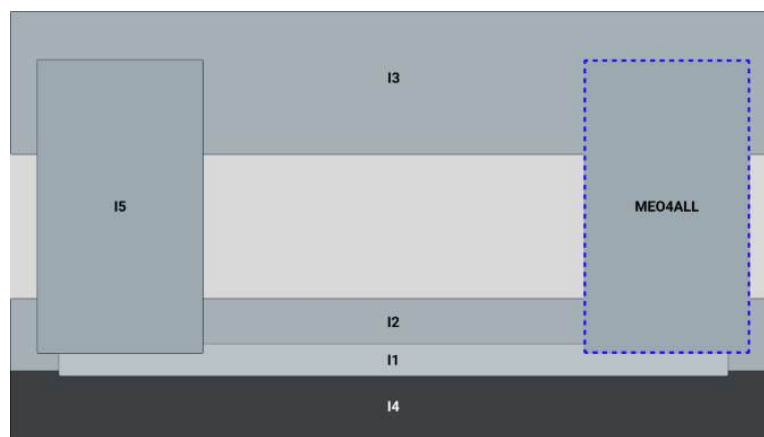


Fig. 2. Final vertical right-side placement of the MEO4ALL interface

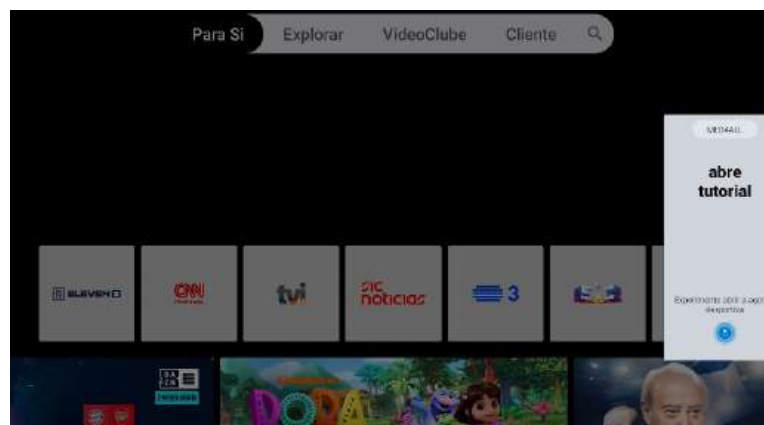


Fig. 3. Final graphical user interface of the MEO4ALL application

4 Discussion

The MEO4ALL redesign demonstrates that accessibility must be embedded in core interface decisions rather than treated as an optional add-on. The vertical right-side placement proved effective in reducing visual clutter while maintaining clear visibility for users who rely on residual vision. For blind users, the spatial separation also helped differentiate the assistant from other system components.

Contextual tips and multimodal feedback supported users with varying levels of digital literacy, bridging gaps between novice and experienced interaction. At the same time, the project highlighted limitations inherent to voice-based systems, such as potential misinterpretation of speech in noisy environments or for users with strong accents. These challenges underscore the importance of complementary interaction modalities.

From a broader perspective, MEO4ALL illustrates how commercial media platforms can integrate AI-driven interaction with inclusive design principles without compromising usability for the general public.

5 Conclusion and Future Work

This paper documented the redesign of the MEO4ALL interface as a case study in accessible iTV design. By combining benchmarking, heuristic evaluation, and technical implementation, the project advanced both practical and theoretical knowledge in accessible media. The new interface improved spatial clarity, feedback transparency, and contextual support for visually impaired users within a real commercial ecosystem.

Future work will extend both the technical and empirical dimensions of this research, offering a more detailed examination of the system's evolution, its validation, and its longer-term impact on accessible interactive television in real context.

Acknowledgments. This study was funded by the Altice Labs@UA project MEO4ALL.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Nevsky, A., Neate, T., Vatavu, R.-D., Simperl, E.: Accessibility research in digital audio-visual media: What has been achieved and what should be done next? In: Proceedings of the ACM International Conference on Interactive Media Experiences (IMX 2023), ACM (2023) <https://doi.org/10.1145/3573381.3596159>
2. Goggin, G., Newell, C.: *Digital disability: The social construction of disability in new media*. Rowman & Littlefield, Lanham (2003)
3. Oliveira, R., Júlio, L., Oliveira, A. P.: MEO4ALL-MEO Android service accessible to visually impaired users. In: Marcus, A., Rosenzweig, E., Soares, M., HCII 2024, LNCS, volume 14714, pp 105–115, Springer (2024). https://doi.org/10.1007/978-3-031-61356-2_7

4. Garcia-Prieto, V., Mora, K., Melero, A.: The accessibility of BBC television to users with disabilities: from the law to user satisfaction. *Profesional de la Información* (2021). <https://doi.org/10.3145/EPI.2021.SEP.18>
5. Deshmukh, A. M.: User experience and usability of voice user interfaces: A systematic literature review. *Information* 15(9), 579 (2024). <https://doi.org/10.3390/info15090579>
6. Costa, D., Duarte, C.: Alternative modalities for visually impaired users to control smart TVs. *Multimedia Tools and Applications* 79, 31931–31955 (2020). <https://doi.org/10.1007/s11042-020-09656-1>
7. Altice Labs: Conversational AI. <https://www.alticelabs.com/products/conversational-ai/> (last accessed 2026/01/12)
8. Dasgupta, R.: *Voice User Interface Design: Moving from GUI to Mixed Modal Interaction*. Apress, Cham (2018)
9. Android Accessibility - TalkBack: <https://support.google.com/accessibility/android/topic/3529932?hl=en-GB> (last accessed 2026/01/12)
10. Google Assistant on TV Platforms. <https://assistant.google.com/platforms/tv/> (last accessed 2026/01/12)
11. Apple Accessibility: VoiceOver on Apple TV. <https://support.apple.com/en-us/105019> (last accessed 2026/01/12)
12. Amazon: Watch videos with Alexa. <https://www.amazon.com/gp/help/customer/display.html?nodeId=G3RVCXTEHPSNJRWW> (last accessed 2026/01/12)
13. Nielsen, J.: 10 usability heuristics for user interface design. Nielsen Norman Group. <https://www.nngroup.com/articles/ten-usability-heuristics/> (1994-2024 updated)
14. Lottie Files: LottieFiles. <https://lottiefiles.com/> (last accessed 2026/01/12)

Open Government, Accessibility, and Limitations to Citizen Participation: The Brazilian Case

João Marcelo Alves Macedo¹[0000-0002-6313-1759] Fábio Henrique Ferreira de Albuquerque²[0000-0001-8877-9634] 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

² Instituto Superior de Contabilidade e Administração de Lisboa – ISCAL, Lisboa, Portugal

Abstract. Growing social engagement in monitoring public spending and demanding transparency expresses citizenship in developed societies and nurtures a culture of open government. Anchored in the Open Government Partnership, open government is defined by transparency, participation, and collaboration between state and society. Yet digital inclusion and meaningful participation still face structural and technological limits linked to accessibility for persons with disabilities. Visual, motor, cognitive, and sensory barriers hinder the use of portals and systems, restricting access to public information and tools for social oversight. In Brazil, the Statute of Persons with Disabilities requires public administration websites to guarantee accessibility in accordance with effective international guidelines. Public pages and applications should also comply with WCAG 2.1 to ensure equal access to information and services. This study applies the WAVE® evaluator and a complementary mobile-focused tool to detect and help correct accessibility barriers on public websites. Evaluating whether people with disabilities can reach and use transparency and open-government portals is essential to gauge the effectiveness, equity, and accountability of these initiatives. Results indicate that, although most pages show some adaptation to mobile screens, persistent problems remain in responsiveness, semantic structure, and usability, often requiring substantial redesign of digital interfaces. These shortcomings reduce trust and widen social and informational inequalities. We conclude that digital inaccessibility undermines the principles of open government: it constrains citizen participation, weakens active transparency, and diminishes mechanisms of social control and accountability. Ultimately, exclusionary design violates the constitutional right of access to information and limits the transformative potential of open-government policies.

Keywords: First Keyword, Second Keyword, Third Keyword.

1 Introduction

In developed societies, citizenship manifests itself through increased social participation in the oversight of public funds and the demand for transparency in government

activities, consolidating itself as a culture of open government. The current model, based on the principles of the Open Government Partnership (OGP), is founded on transparency, participation, and collaboration between the state and civil society, with the aim of strengthening public trust and promoting the co-production of public policies.

The OGP works to ensure that member countries, such as Brazil, one of the founding countries since 2011, implement National Action Plans that include concrete commitments on issues such as active transparency, access to information, and social participation (OGP, SD) [1]. These commitments generally lead to the creation of specific content and functionalities on government portals, namely: (a) organization and improvement of public transparency portals, aiming at more accessible and interoperable data, in accordance with the principles of open data and social involvement (Ministry of Management and Innovation in Public Services, 2025) [2]; (b) commitments aimed at increasing the usability and access to databases on the Transparency Portal and other official platforms, expanding their reach and effectiveness for citizens (OGP, SD) [1].

These commitments are directly reflected in demands or objectives that influence the content made available on public websites, which contributes to making government information more systematic, readable, and useful for society as a whole.

In the Brazilian context, the 2022 Demographic Census projected that, among the 14.4 million individuals with disabilities in Brazil, approximately 7.9 million face vision difficulties, even when using glasses or contact lenses, which characterizes the most common functional limitation among people with disabilities in the country [3]. The situation is alarming due to the lack of appropriate treatment, considering that the World Health Organization (WHO) indicates that approximately 285 million individuals worldwide have vision problems, with most cases, between 60% and 80%, being preventable or having treatment alternatives. Given this context, concerns arise about the influence of government websites on this audience.

On the other hand, effective digital inclusion and social participation face structural and technological limitations associated with accessibility for people with disabilities. It appears that transparency and digital government are more restrictive than inclusive. Visual, motor, cognitive, or sensory obstacles that hinder the use of open government portals and systems have been documented, limiting access to public information and interaction with transparency and social oversight tools.

In the Brazilian context, Law No. 13,146/2015 (Statute of Persons with Disabilities) [5] establishes, in its Article 63, that public administration websites must ensure accessibility for individuals with disabilities, in accordance with international best practices and guidelines. Additionally, Decree No. 10,645 of 2021 [6] establishes that all public pages and applications must comply with digital accessibility standards, with special attention to the Web Content Accessibility Guidelines (WCAG 2.1), aiming to ensure equal access to information. Despite these standards, it is observed that many governments neglect to systematically verify accessibility on their portals and applications. Recently, efforts have been focused on Brazilian Sign Language (LIBRAS). One of the initiatives aimed at specific audiences was VLibras, aimed at people with hearing impairments.

Macêdo, Becker, and Sá (2023) [7] had already demonstrated that, for all municipalities in the state of Paraíba, comparing access via web browsers and mobile devices, according to an experimental strategy, shows how responsiveness works on the public transparency websites of city halls in Paraíba. The authors concluded that most websites are responsive, but where this compatibility is lacking, the dissemination of access to public information is negatively affected.

Macedo et al. (2025) [8] conclude that the existence of portals alone does not guarantee effectiveness. Poor usability and lack of accessibility act as barriers that facilitate non-accountability for managers. Many managers adopt a reactive stance, complying with the law only to avoid punishment, without a genuine interest in promoting transparency through quality interfaces. Metaphorically, the findings indicate that the transparency portals of Brazilian capitals are like public buildings with access ramps that are too steep and confusing signage. Although the building “exists” and most people (85%) believe they can enter it (perception), in practice, when they try to navigate the corridors (mobile usability and memory load) or if they have specific needs (accessibility), they encounter locked doors or paths that lead nowhere (Macêdo et al., 2025) [8].

One way forward would be to improve accessibility through web interface evaluation tools. According to the Institute for Disability Research, Policy & Practice (2024) [9], Web Accessibility Evaluation (WAVE) is a web browser extension that helps identify accessibility issues and provides recommendations for improving the experience of users with disabilities. This tool presents a list of several common accessibility challenges. The WAVE extension was configured according to the parameters necessary for accessibility analysis.

The WAVE extension works by evaluating the page structure, identifying possible violations of accessibility standards, and flagging them with suggestions for corrections. In this sense, various aspects were considered, such as HTML semantics, attributes, ARIA (Accessible Rich Internet Applications), color contrast, readability, keyboard navigation, and others. For the purposes of this research, the following issues were highlighted: errors, contrast problems, resources, alerts, and ARIA. When examining websites, especially analyzing their accessible interface in terms of accessibility, it is essential to ensure equal access to information, regardless of physical disabilities and other possible limitations. In the study by Ahmi & Mohamad (2016) [11], AChecker and WAVE are used to assess accessibility. The results indicate a low level of compliance with the guidelines established in WCAG 2.0 and Section 508 (Ahmi & Mohamad, 2016) [11].

Among the points that should be considered are the provision of text alternatives for any non-textual content, keyboard accessibility, and color contrast (Ahmi & Mohamad, 2016) [11]. In addition, other issues such as adaptability, input assistance, compatibility, empty links, and empty headers can be improved.

However, we still have not adopted an integrated policy that addresses the other dimensions of disability. Given this context, our research is designed to evaluate state government websites, which should be built based on these guidelines. In this study, we use the Web Accessibility Evaluation Tool (WAVE®) and another tool for mobile devices, which helps identify and correct accessibility barriers in open government

website content on the web. Analyzing the ability of individuals with disabilities to access and use transparency and open government portals is an essential step in quantifying the effectiveness, equity, and accountability of open government initiatives.

2 Theoretical framework

The efficiency of public data provision constitutes a central element in the realization of governmental transparency, as it is intrinsically dependent on citizens' capacity to access, understand, and interact with the information systems made available by the state. This perspective aligns with the principles of the Open Government Partnership (OGP), which conceives open government not merely as the formal release of data but as a model grounded in usability, reusability, and the potential of data to foster civic participation, social oversight, and accountability (OGP, n.d.). [1]. Accordingly, the mere publication of information does not guarantee substantive transparency if data are not presented in an intelligible, accessible, and user-centered manner.

In the Brazilian context, Law No. 12,527/2011 (Access to Information Law - LAI) [12] establishes access to public information as the rule and secrecy as the exception, imposing upon public administration the duty to promote active transparency through the clear, objective, and comprehensible disclosure of information of collective interest. However, the effectiveness of this fundamental right is conditional upon the existence of digital interfaces that enable the full use of information; otherwise, access remains merely formal. Recent empirical evidence indicates that, for open data to achieve its democratic potential, users must perceive such data not only as accessible but as effectively usable, allowing for analysis, interpretation, and reuse in decision-making and public oversight processes (Macedo et al., 2025) [8].

In this regard, the Web Content Accessibility Guidelines (WCAG 2.1) [13] emphasize that the provision of public information must comply with the principles of perceivability, operability, understandability, and robustness, ensuring that persons with disabilities can access and use digital content on an equal basis. Failure to comply with these standards undermines both the objectives of the OGP and the substantive implementation of the LAI by creating informational barriers that restrict civic participation and exacerbate structural inequalities in access to public information. Consequently, efficiency in public data provision requires the integration of open government principles, digital accessibility, and user-centered design, under penalty of weakening the mechanisms of social control and accountability that underpin contemporary democracies.

Moreover, the Organisation for Economic Co-operation and Development (OECD), in its analysis of the evolution of digital government solutions, identified that advancements in public services are closely associated with the expansion of wireless connectivity infrastructures, which enhance access opportunities through mobile devices (OECD, 2011) [14]. This shift accelerated the transition from electronic government (e-government) to mobile government (m-government), whereby access to public services increasingly occurs via applications rather than exclusively through traditional web browsers. This transformation reinforces the need to assess not only web portals

but also mobile interfaces, considering their specific requirements regarding usability, accessibility, and interaction.

Within this context, overcoming the comprehension barrier becomes imperative. Usability can be defined as the extent to which a system can be used effectively, efficiently, and satisfactorily, enabling citizens to recognize the value of the information provided. Interfaces that are confusing or impose excessive cognitive load compromise the understanding of budgetary and governmental data, thereby undermining the objectives of public transparency. The communicative process, therefore, does not end with the publication of a portal; it depends fundamentally on the quality of human–computer interaction (HCI). While the intrinsic value of public data may be indisputable, their quality from an interaction standpoint is entirely dependent on users' ability to operate the system adequately.

HCI extends beyond usability to encompass accessibility, user experience, and social interaction. The adoption of user-centered design techniques from the early stages of developing civic participation platforms can significantly contribute to the creation of more intuitive, inclusive, and accessible interfaces, particularly for users with varying levels of ability and technological experience (Macedo et al., 2025) [8].

The evaluator will select the method that best suits their evaluation needs and available resources. IHC evaluation methods can be divided into: investigations, user observations, and inspection (Barbosa et al., 2021) [10].

Within this methodological framework, the Web Accessibility Evaluation Tool (WAVE) emerges as a relevant instrument for assessing the accessibility of governmental websites. Developed by WebAIM, the tool supports researchers, designers, and developers in systematically identifying accessibility barriers, providing detailed diagnostics, and providing corrective recommendations. WAVE categorizes its findings into six analytical dimensions, errors, contrast errors, alerts, features, structural elements, and ARIA attributes, enabling a comprehensive evaluation of access conditions for users of assistive technologies.

In the Brazilian context, Law No. 13,146/2015 (Statute of Persons with Disabilities) [5] establishes, in Article 63, that public administration websites must ensure full accessibility for individuals with disabilities, in accordance with international best practices and guidelines. This legal provision recognizes access to public information as a fundamental right of an instrumental nature, whose effectiveness depends on the elimination of technological and communicational barriers. Complementarily, Decree No. 10,645/2021 [6] mandates that all public websites and applications comply with digital accessibility standards, with particular emphasis on the Web Content Accessibility Guidelines (WCAG 2.1). Thereby consolidating accessibility as a structural requirement of digital transformation in the public sector and as a necessary condition for informed civic participation.

Despite this robust normative framework, a significant number of governments continue to neglect the systematic assessment of accessibility in their institutional portals and applications, often treating legal compliance as merely formal or declaratory. The absence of continuous technical evaluation and monitoring processes compromises the intelligibility and usability of public information, disproportionately restricting access for individuals with visual, motor, or cognitive disabilities. From the perspective of

accountability as conceptualized by Campos (1990) [15], such informational limitations weaken society's capacity to monitor, evaluate, and question state action. Since accountability can only materialize when citizens have access to information that is not only available but also accessible and comprehensible.

Recently, governmental efforts in digital accessibility have largely focused on Brazilian Sign Language (LIBRAS), with initiatives such as VLibras aimed at people with hearing impairments. While these initiatives represent important advances, their isolated implementation reveals a fragmented approach to accessibility, insufficient to address the multiple dimensions of disability in an integrated manner. This fragmentation reproduces informational inequalities and undermines substantive accountability, insofar as, according to Campos (1990) [15], social control is not exhausted by the existence of formal transparency mechanisms but depends on the social capacity to demand, understand, and effectively use public accounts. Consequently, the absence of a comprehensive digital accessibility policy aligned with international standards such as WCAG 2.1 limits the democratic potential of open government and weakens the foundations of transparency, participation, and public accountability.

3 Methodology

This study adopts a deductive methodological approach, of a theoretical–empirical nature, aimed at the systematic analysis of digital accessibility in web and mobile portals of public institutions. This approach enables the articulation of normative and conceptual foundations—particularly those related to open government, accessibility, and accountability—with empirical evidence derived from the direct evaluation of governmental digital interfaces.

For the purposes of the analysis, official government websites of the 26 Brazilian states, the Federal District, and the federal government were selected, as they represent central institutional references for the dissemination of public information and the promotion of transparency within their respective jurisdictions, totaling 27 federative units and one Federal District. Data collection was conducted on 31 October 2025, considering exclusively the information available at the time of access, with particular emphasis on each institution's homepage, as it constitutes the primary entry point for citizens seeking institutional content and digital public services.

Interface responsiveness was examined through the assessment of portal behavior during the adaptation process to different screen sizes, focusing on the appropriate adjustment of visual and functional elements. All evaluations were conducted in a desktop environment, using updated web browsers, in order to ensure technical consistency and comparability across the analyzed portals.

Accessibility assessment was carried out using the WAVE Evaluation Tool browser extension, a widely recognized instrument in both academic literature and professional practice for identifying accessibility barriers in web environments. Developed by WebAIM, the tool supports the detection of common accessibility issues and provides technical recommendations for remediation, thereby contributing to the improvement of user experience for individuals with disabilities. Prior to the evaluation process, the

WAVE extension was carefully configured according to standardized parameters to ensure the reliability and consistency of the analyses.

WAVE operates by analyzing the structural and visual composition of web pages, identifying potential violations of accessibility standards and highlighting them through visual indicators accompanied by interpretative guidance. In this process, critical aspects such as HTML semantic structure, appropriate use of Accessible Rich Internet Applications (ARIA) attributes, color contrast, text legibility, keyboard navigation, and content organization are assessed. For the purposes of this study, the analysis focused on the following key categories: errors, contrast errors, features, alerts, and ARIA attributes, as these dimensions constitute essential indicators of accessibility conditions for users who rely on assistive technologies.

4 Presentation and discussion of data

The data generated by the WAVE Evaluation Tool were systematically tabulated and organized, and subsequently classified according to the specific criteria established by the WCAG 2.2 (Web Content Accessibility Guidelines). Under these guidelines, websites are required to meet at least Level A conformance to ensure that persons with disabilities can effectively access and use digital content. This methodological framework enabled the assessment not only of the presence of technical accessibility barriers, but also of the degree to which governmental portals adhere to internationally recognized digital accessibility standards.

The WCAG framework is grounded in four foundational principles for effective accessibility: perceivability, operability, understandability, and robustness. These principles guide the evaluation of how information is presented to users, the extent to which content can be interacted with through different input modalities (including keyboard navigation and assistive technologies), the clarity of structure and language, and the compatibility of digital systems with a wide range of user agents and screen readers. Compliance with these principles is essential to ensuring equitable access to public information and enabling meaningful civic participation in digital government environments.

Empirical findings revealed substantial heterogeneity in accessibility performance across the analyzed portals. The website of the State of Piauí achieved the highest overall performance, with an AIM Score of 9.9, indicating a high level of conformity with the criteria assessed by WAVE. Conversely, the State of Rondônia recorded the highest number of contrast-related issues, totaling 122 occurrences, followed by the Federal District with 103 issues, highlighting significant shortcomings with respect to the principle of perceivability. The portal of the State of Rio de Janeiro exhibited the highest number of alerts, with 179 instances, signaling potential accessibility issues that require manual verification and structural adjustments. Additionally, the Federal District presented 408 ARIA elements, which may reflect either an effort to enhance accessibility or the excessive or improper use of ARIA attributes, thereby necessitating careful scrutiny regarding their correct implementation.

States	State abbreviations	Regions	General Diagnosis of the SITE						Frequency in the Region	Frequency in Brazil
			Errors	Contrast Errors	Alerts	Features	Structure	ARIA		
Goiás	GO	Midwest	4	4	80	61	97	90	7,7	1,0
Mato Grosso	MT	Midwest	10	22	9	7	26	89	4,7	2,0
Mato Grosso do Sul	MS	Midwest	62	10	7	4	1	12	2,3	3,0
Distrito Federal	DF	Midwest	81	103	132	25	82	408	2	4,0
Piauí	PI	Northeast	0	1	3	1	9	64	9,9	1,0
Bahia	BA	Northeast	8	1	91	14	39	151	6,8	2,0
Rio Grande do Norte	RN	Northeast	6	10	26	9	63	18	6,2	3,0
Pernambuco	PE	Northeast	25	12	19	12	182	256	5,4	4,0
Ceará	CE	Northeast	14	11	30	11	50	21	5	5,0
Alagoas	AL	Northeast	27	19	11	3	45	54	4,3	6,0
Paraíba	PB	Northeast	35	5	38	12	52	1	4,3	7,0
Maranhão	MA	Northeast	59	9	76	98	15	135	3,6	8,0
Sergipe	SE	Northeast	23	28	53	37	5	49	3	9,0
Tocantins	TO	North	3	0	32	1	2	173	8,9	1,0
Acre	AC	North	10	21	3	2	25	23	5,1	3,0
Amazonas	AM	North	9	16	21	46	15	38	5,1	2,0
Amapá	AP	North	6	25	23	7	43	35	4,7	4,0
Roraima	RR	North	5	26	64	43	46	103	4,5	5,0
Pará	PA	North	55	0	21	14	37	175	3	6,0
Rondonia	RO	North	22	122	27	62	59	62	1,4	7,0
Espirito Santo	ES	Southeast	1	25	131	20	70	95	5,3	1,0
São Paulo	SP	Southeast	31	1	41	12	43	147	4,9	2,0
Minas Gerais	MG	Southeast	27	21	16	54	43	270	3,6	3,0
Rio de Janeiro	RJ	Southeast	303	8	179	15	39	11	1	4,0
Rio Grande do Sul	RS	South	3	0	58	31	14	18	8,4	1,0
Santa Catarina	SC	South	11	11	18	7	48	23	6	2,0
Paraná	PR	South	90	36	17	17	142	8	1,8	3,0

Table 1: Wave assessment of Brazilian state transparency websites
Font: research data (2025)

The findings demonstrate that, although most of the analyzed portals exhibit some degree of adaptation to mobile device screens, significant structural deficiencies persist in terms of responsiveness, semantic content organization, and interface usability. Such limitations undermine effective user interaction with digital systems and, in several cases, indicate the need for a comprehensive redesign of digital interfaces, rather than merely incremental or superficial adjustments. These results underscore that partial technological adaptation is insufficient to ensure substantive accessibility and digital inclusion.

It is therefore concluded that digital inaccessibility critically undermines the principles of Open Government, as it restricts civic participation, limits the exercise of social oversight, and compromises the effectiveness of active transparency, as established under Law No. 12,527/2011 (Access to Information Law). By impeding equitable access to public information, accessibility barriers weaken mechanisms of accountability, thereby violating the constitutional right to access information enshrined in Article 5, item XXXIII, of the 1988 Federal Constitution.

In this regard, ensuring comprehensive, continuous, and standards-based digital accessibility must be understood as a structural prerequisite of contemporary democracy, rather than as a secondary technical requirement. The development of accessible governmental digital environments is indispensable for the effective democratization of the State, the strengthening of public trust, and the full inclusion of all citizens in the public sphere—particularly those historically marginalized from political and informational participation.

References

1. Open Government Partnership. (n.d.). Restructuring of the Brazilian Federal Government Transparency Portal (BR0067). <https://www.opengovpartnership.org/members/brazil/commitments/BR0067>
2. Ministério da Gestão e da Inovação em Serviços Públicos. (2025). Políticas e planos de ação para governo aberto. <https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/estrategias-e-politicas-digitais/politicas-e-planos-de-acao-para-governo-aberto>
3. Instituto Brasileiro de Geografia e Estatística. (n.d.). Censo 2022: Brasil tem 14,4 milhões de pessoas com deficiência. <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/43463-censo-2022-brasil-tem-14-4-milhoes-de-pessoas-com-deficiencia>
4. Ministério da Saúde. (2023). OMS alerta que 2,85 bilhões de pessoas no mundo têm a visão prejudicada. <https://www.gov.br/saude/pt-br/assuntos/noticias/2023/fevereiro/oms-alerta-que-285-milhoes-de-pessoas-no-mundo-tem-a-visao-prejudicada>
5. Brasil. (2015). Lei nº 13.146, de 6 de julho de 2015: Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Diário Oficial da União.
6. Brasil. (2021). Decreto nº 10.645, de 11 de março de 2021: Regulamenta o art. 75 da Lei nº 13.146, de 6 de julho de 2015, para dispor sobre as diretrizes, os objetivos e os eixos do Plano Nacional de Tecnologia Assistiva. Diário Oficial da União.
7. Macêdo, J. M. A., Becker, V., & de Sá, F. M. F. (2023, July). Heuristic-Based Evaluation of Transparency Websites of the Municipal Governments Viewed on Web and Mobile Browsers. In *International Conference on Human-Computer Interaction* (pp. 434-454). Cham: Springer Nature Switzerland. Homepage: https://link.springer.com/chapter/10.1007/978-3-031-35599-8_29

8. Macedo, J. M. A., Becker, V., Feliciano de Sá, F. M., Rocha Filho, E. V. V. da, & Cavalcanti, D. de Q. (2025). Avaliação de sistemas de Governo Aberto e de transparência pública nas capitais brasileiras. *Cadernos Gestão Pública E Cidadania*, 30, e90832. <https://doi.org/10.12660/cgpc.v30.90832>
9. Institute for Disability Research, Policy & Practice (2024). Web Accessibility Evaluation tool [Internet]. Logan, UT; 2024 [cited 2024 Jan 30]. Available from: <https://wave.webaim.org/>.
10. Barbosa, S. D. J., Silva, B. D., Silveira, M. S., Gasparini, I., Darin, T., & Barbosa, G. D. J. (2021). Interação humano-computador e experiência do usuário. Auto publicação.
11. Ahmi, A. and Mohamad, R., Evaluating Accessibility of Malaysian Public Universities Websites using AChecker and WAVE (December 1, 2016). *Journal of Information and Communication Technology*, 2016, Available at SSRN: <https://ssrn.com/abstract=3550314>
12. Brasil. (2011). Lei nº 12.527, de 18 de novembro de 2011: Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. *Diário Oficial da União*.
13. World Wide Web Consortium. (2018). Diretrizes de acessibilidade para conteúdo web (WCAG) 2.1. <https://www.w3.org/TR/WCAG21/>
14. Organisation for Economic Co-operation and Development. M-government: mobile technologies for responsive governments and connected societies. Paris:OECD Publishing. 2011
15. Campos, A. M. (1990). Accountability: quando poderemos traduzi-la para o português?. *Revista De Administração Pública*, 24(2), 30 a 50. Recuperado de <https://periodicos.fgv.br/rap/article/view/9049>

EEG and Affective Computing



Dataset characterization for user emotional analysis

Joan Josep Ordoñez¹[0009-0002-7536-1874], Jose Maria Buades¹[0000-0002-6137-9558], Francisco José Perales¹[0000-0002-9872-3172], Sandra Baldassarri²[0000-0002-9315-6391], José Ramón Beltrán²[0000-0002-7500-4650], and Gustavo Javier Astudillo³[0000-0003-0645-6335]

¹ Universitat de les Illes Balears, 07122 Palma, Islas Balears, España
joanjosep.ordonez@uib.es

² Universidad de Zaragoza, 50009 Zaragoza, España

³ Universidad Nacional de La Pampa, Santa Rosa, Argentina

Abstract. The study of emotions from human behavior is the core part of affective computing and plays a key role on Human-Computer Interaction. In the existing literature multiple datasets can be found compiling information about participants in different contexts reacting to emotional stimuli. In this work we present an analysis of a total of 84 emotional datasets published between 2014 and 2024, found in multiple repositories such as Springer, IEEE Xplore, ACM and Nature. This study offers a general characterization of 84 datasets used in affective computing, including channels used, labeling models and description, and volume of data. With this work, we find preference in the use of multimedia modalities for emotion recognition, as well as a general lack of well-balanced data regarding age and gender.

Keywords: Datasets · Affective computing · Emotional analysis

1 Introduction

With the increase in the number of systems that aim to integrate affective user analysis in the behavior of interactive systems, the study of emotions is a growing field from a computational standpoint. Emotion recognition (ER) is a field in which many systems have already been developed, from robot companions for child patients during stressful clinical procedures [12], to interactive TV systems implementing emotion recognition [18], this field promises a more human-centered design in developing technologies.

As affective computing focuses in the use of AI technology to analyze emotional expressions in both single and multi-channel modalities, there is a basic component that cannot be ignored, that is, the data available. Emotional datasets are a type of dataset that focuses on collecting emotional responses from users on various channels and labeling them with regards to the expressed emotion, either expected from the stimuli or felt from that same participant [1].

With this raw data, emotional datasets are a valuable asset in affective computing. However, depending on the type of application (see Figure 1), the available data may not suffice for training an AI model, and thus, many studies rely on self-sourcing the data, that is, building the dataset from scratch by experimental design that meets their specific needs [7, 20]. This tendency can become a problem, as the data collected is not always easily available for the research community for further contributions. In order to help in this field, we carried out this study to summarize general tendencies on the available datasets to guide further efforts in manufacturing new datasets. As opposed to other reviews regarding multimodal emotion recognition, this work focuses specifically on finding common dataset definition trends, as well as highlighting key aspects for improvements in the general corpus of available datasets.

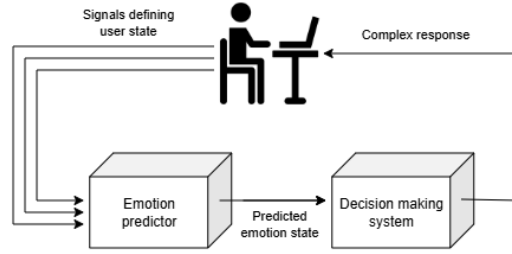


Fig. 1. Generic example of affective computing in HCI. The user's signals are used to predict its emotional state, which is then used in a more complex decision making system, i.e. TV recommendation systems or intelligent agent responses.

This paper is structured as follows: section 2 shows the methodology and research questions followed to obtain and analyze the datasets, section 3 presents the results and discussion, and section 4 gives the conclusions of this work.

2 Methodology

This study aims to give a broad perspective on common trends regarding available emotional datasets. To achieve this, we performed a review on existing publicly available datasets with the goal of answering the following questions: 1) From which channels is emotional information sampled? 2) How is the labeling performed? Both from a theoretical model perspective and from experimental execution 3) What are the most common bias trends in these data sources? and 4) What data is publicly accessible for research purposes?

To perform this study, we aimed at analyzing available open datasets. To this end, we started from the already existing set of literature reviews. We selected 6 literature reviews from the result of searching "Multimodal Emotion Recognition" + "systematic review" in Google Scholar in the period of 2020 to 2024

that included explicit references to emotional datasets. The reviews chosen with these criteria are [5, 6, 8, 15, 16, 21].

From the selected reviews, we obtained all the available references to datasets, adding up to a total of 98 papers. Through the analysis of these references, we further added 74 references to other datasets, excluding duplicates. Thus, 172 references were obtained. The sole criterion for dataset eligibility was the availability of a reference (e.g., a paper or repository link) describing the data corpus. Accordingly, the selected sources include journal publications, conference proceedings, workshops, and links to Internet repositories.

In order to filter which datasets fit the purposes of this study, we selected the following as inclusion criteria: 1) the document describing the dataset is accessible, 2) the publication date is within the range 2014-2024, ends included, and 3) the dataset must include labels regarding the emotional expression of the participants, that is, either categorical emotions (i.e. happy, sad), or dimensional representations (i.e. valence-arousal). Note that dimensional approaches do not correspond 1 to 1 to any emotion and may be used to explore other facets of human psychology, such as stress, sleep quality [13] or depression [19]. With these criteria, from the 172 eligible publications, we selected a total of 84 datasets. The references and characterization are available as additional resources.

3 Results

3.1 From which channels is emotional information sampled?

To source information regarding emotional expressions, it is necessary to define the modalities from which to collect the data. There are datasets that use an unimodal approach, like FER-Wild [11] or SEED [22], and others in which a multimodal approach is chosen, like RAVDESS [9] or ASCERTAIN [17]. Regardless of the approach, we analyzed the modalities used, as shown in Figure 2.

As can be extracted from the results, the most common modalities are those related to multimedia content, since these are the most easy modalities to collect and process. There are some signals, such as video and movements, that can be further divided into specific experimental setups, such as 3D vs thermal video, or head vs eye movements, but have been combined for simplicity sake. Regarding purely physiological signals, it can be deduced that the most common modalities are electroencephalogram (EEG), galvanic skin response / electro-dermal activity (GSR / EDA), and electrocardiogram (ECG). Note how there are 7 signals related to heart and blood activity included in the signals, from impedance cardiography (ICG) to hemodynamic response (Hemo) to oxygen saturation in blood (SpO2). This shows how there is clear interest in exploring the relation between heart and blood activity and emotional responses.

By deeper understanding of the signals, some signals can be extracted from others with processing methods, such as inter-beat intervals IBI from ECG. The specifics of including these types of signals come from the experimental setup, in which the sensors used return the already processed signal and not the raw

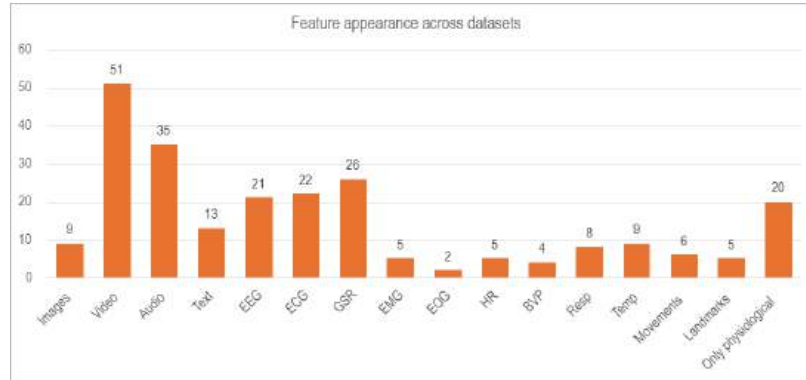


Fig. 2. Unique signals used in the studied datasets. Some signals may appear in different formats in the same dataset (i.e. RGB and infrared video) and are counted once.

components. This also happens in some EEG datasets, where instead of giving the raw node outputs, the processed components are given [17].

3.2 How is the labeling performed?

As a relevant part of the emotional datasets, labels give the association between the sampled data and the corresponding emotion of the participant. However, there are many ways in which the labeling process can be performed. Thus, we analyzed it by dividing it into three key components: theoretical model used, dimensionality of the labels, and perspective of the labeler. Note that in the results given, some datasets use different approaches at the same time, since the dataset is formed as a combination of multiple experiments in different conditions.

From all the available emotion models, we performed a division in two main groups, like in [2]: categorical models, in which the emotions are expressed as discrete labels by name, like the 6 emotion model by Ekman [4], and dimensional models, in which emotions are expressed along a set of continuous dimensions, like the valence-arousal model [14]. We found that categorical models are the most common approach, with 60 datasets using them. However, not all models are compatible, as different emotion names may be used in different sets. The next most common model used is the 2D emotion model, with 27 different datasets using it. Lastly, only 8 datasets use a 3D emotion model.

With a model chosen, the following aspect to consider is the scale in which the emotions are represented (i.e. discrete, continuous). In this manner, categorical emotions can use binary scales to show presence or absence of an emotion or discrete scales to represent emotion intensity. On dimensional emotions, either a discrete or a continuous scale can be used to obtain different granularity in the data. From the datasets analyzed, we found a balance between binary and discrete scales, with 41 and 42 datasets respectively, and only 11 datasets using

a continuous scale approach. Note how, despite being less commonly used, from continuous scales a discrete representation can be obtained, and not otherwise.

Lastly, the perspective from which the labeling is performed is relevant to the resulting dataset, since it is not equivalent the emotion felt by a person than the emotion perceived by another from the expressions given, as the expression may be faked or the person may not be able to fully express the real emotion. With this idea, we found that only 34 datasets out of 84 include labels from the perspective of the experiment subject, and a total of 43 datasets performed the labeling with external subjects. As a third perspective, 13 datasets prefer to approach this problem by already defining the label prior to the experiment, that is, by association of the stimuli to an emotion, that same emotion labels the experiment, since it is the expected expressed emotion.

3.3 What are the most common bias trends in these data sources?

As dataset manufacturing is a complex process, there are many aspects that need to be considered regarding the bias in the results. This bias mostly comes from the participants, as they constitute the sample population of the dataset. The most common biases in human-related dataset manufacturing refer to age and gender but also can relate to the spoken language of the participants. By this lack of data, population-aimed applications on the lower represented groups require further development work, as model development and performance is constrained to the available data.

Analyzing the datasets, it was found that only 49 and 50 gave information about the age and gender distribution of the subjects, respectively. The age distributions given have been split into four groups, as shown in Table 1. This table clearly shows that almost all datasets include samples in the age group between 18 and 40, yet note the lack of samples in the remaining age groups, especially the extremes. This shows a clear lack of emotional datasets representing the population in these groups. Regarding gender, it was found that only 18 datasets have male-female subject balance, and the remaining 32 are unbalanced.

Table 1. Age group distribution. Counts include datasets with samples that fall in the defined ranges but are not exclusive. (n=49)

Age group	Below 18	Between 18-40	Between 40-65	Over 65
Count	8	48	19	11

Another relevant bias regarding audio and text related datasets is the language used, as it constitutes a non-negligible part of human communication. The languages used are summarized in Table 2. It can be seen that the most common language is English (EN), followed by Chinese (ZH). The remaining counts show that there is interest in studying emotion recognition using features from multiple other languages, yet public datasets are lacking.

Table 2. Spoken languages in the datasets (ISO 639 codes). Some datasets include multiple languages. (n=29)

Language	EN	ZH	TR	ZH(Mand)	SV	HE	RU	DE	EL	HU	SR
Count	18	6	3	2	1	1	1	1	1	1	1

3.4 What data is publicly accessible for research purposes?

The last research question aims to give insight into the availability of the datasets. Since all 84 selected datasets have explicit references, we studied the ways the files could be retrieved. It was found that a total of 47 datasets have a working link to a web page where the files can be either retrieved or requested, 15 datasets had links that could not be accessed, either because of regional restrictions or are no longer available, and lastly the remaining 22 datasets do not have any explicit link. These results show some clear problems regarding data availability. The discontinuation of the distribution of datasets causes the corpus of available data to slowly shrink over time, losing valuable and potentially unique sources of emotional data. Moreover, the generation of large datasets forces some authors to not be able to publish the data on the Internet, as seen from older datasets like SAMM [3], and thus adding complexity to the retrieval of these resources.

Another perspective on availability is the volume of available samples from each source. To this end, we studied the unique number of participants on each dataset, when available, and analyzed the nature of the dataset itself. This last component regards the use of original sourced data from an experimental setup or the use of large amounts of data obtained from external sources with techniques like web-scraping. These techniques do not ensure the uniqueness of each participant as typically person recognition is not used to remove potential duplicates (i.e. AffectNet [10]). With this consideration, the resulting numbers of unique participants from each modality, considering that no participant is included in two different datasets, are summarized in Table 3.

Table 3. Unique subjects for each of the modalities used. A distinction between external sources included and not included in the resulting counts.

Signal	Audio	Images	Video	Text
Unique subj	5795	474250	16134	2252
Unique subj (no external)	1934	250	2512	689

Table 3 shows how the use of external sources is most common in multimedia related modalities. Note how the biggest impacted modality is images, as the largest datasets use images as the preferred modality in facial ER. From the remaining modalities, no external sourcing methods are used, as shown in Figure 3. These counts also show a clear relation between the most used modalities and the resulting available data. However, a lack of available data is exposed for EEG, despite being one of the most used modalities in ER.

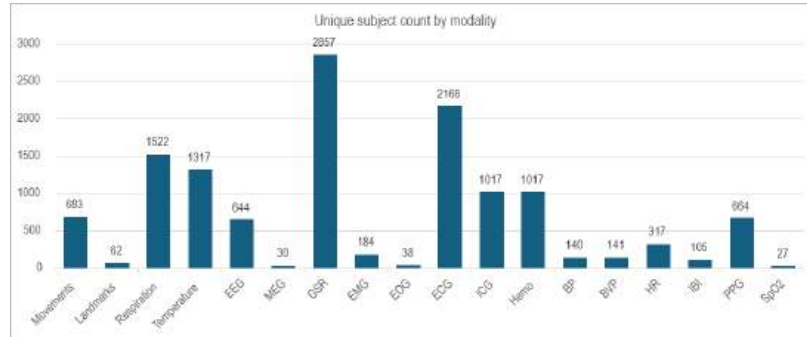


Fig. 3. Unique subjects for each of the modalities used without external sources.

3.5 Proposed guidelines

As a means to aid on new dataset generation, we highlight a set of recommendations based on the evaluations obtained in this work. (1) Source selection should be focused on using general signals whenever possible, that is, to include signals that can be further analyzed by decomposition or cleansing, such as raw ECG readings instead of HR readings. (2) As different labeling modalities exist, it should be very clearly defined the emotion model used, and the perspective from where the emotional label is obtained. (3) Sample selection and description regarding age, gender and ethnicity, and speech when collected, is a must when defining a dataset, as this can help use more fairly distributed data during model training. And (4) by knowing data availability, it is significant to attempt to contribute on modalities with data scarcity, as by relying only on a few datasets for comparing methods does not ensure complete generalization of models.

4 Conclusions

In this work, the main features of 84 emotional datasets have been reviewed and discussed, providing a more detailed characterization of the data available for emotion recognition. Based on the results obtained, the most used sources are audio and video, but also EEG, ECG and GSR. This shows a lack of available data for applications that focus on 3D spatial information. Moreover, the preferred method to represent emotions is with a categorical approach, as it is easily understandable by both participants and researchers, as exposed in various datasets studied. The use of these representations, however, pose a problem regarding compatibility, as the emotions used may not be fully compatible from a strictly psychological point of view (i.e. joy and happy). The main potential biases have also been studied, since the inclusion of biased data can impact on the generalization capabilities of the resulting trained models. Regarding this aspect, it was found that the main age group included is between 18 and 40, lacking

datasets in other age groups, and that only 21% of the datasets showed a balanced gender distribution, thus applications may work better on specific genders and age groups. Lastly, we also studied dataset availability, finding 15 datasets that are no longer available despite being used in current research efforts.

With these results, we aim to give insight into the main gaps in available data and guide the process of dataset selection for emotion recognition applications.

Acknowledgments. This work was possible thanks to the support from the projects PLEISAR (Playful Experiences with Interactive Social Agents and Robots) PID2022-136779OB-C31/C32 (MINECO/AEI/ERDF, EU), and TEMOR, TED2021-130374BC22 (with funding from MCIN/ AEI/10.13039/ 501100011033/ and European Union NextGenerationEU/ PRTR), and with the co-funding of the Aragonese Government (Group T60_23R).

Disclosure of Interests. The authors do not show any conflict of interests related with the content of this article.

References

1. <http://kahlan.eps.surrey.ac.uk/savee/Introduction.html>
2. Astudillo, G.J., Sanz, C.V., Baldassarri, S.: Emotionally-tagged video database: a systematic review. *Data Technologies and Applications* pp. 1–26 (11 2025). <https://doi.org/10.1108/DTA-06-2024-0664>
3. Davison, A.K., Lansley, C., Costen, N., Tan, K., Yap, M.H.: Samm: A spontaneous micro-facial movement dataset. *IEEE Transactions on Affective Computing* **9**(1), 116–129 (2018). <https://doi.org/10.1109/TAFFC.2016.2573832>
4. Ekman, P.: Facial expressions of emotion: New findings, new questions. *Psychological Science* **3**, 34–38 (1992). <https://doi.org/10.1111/j.1467-9280.1992.tb00253.x>
5. Gahlan, N., Sethia, D.: Federated learning in emotion recognition systems based on physiological signals for privacy preservation: a review. *Multimedia Tools and Applications* (2024). <https://doi.org/10.1007/s11042-024-19467-3>
6. Geetha, A.V., Mala, T., Priyanka, D., Uma, E.: Multimodal emotion recognition with deep learning: Advancements, challenges, and future directions. *Information Fusion* **105** (5 2024). <https://doi.org/10.1016/j.inffus.2023.102218>
7. Hasnul, M.A., Ab. Aziz, N.A., Abd. Aziz, A.: Augmenting ecg data with multiple filters for a better emotion recognition system. *Arabian Journal for Science and Engineering* **48**(8), 10313–10334 (Aug 2023). <https://doi.org/10.1007/s13369-022-07585-9>
8. Kalateh, S., Estrada-Jimenez, L.A., Nikghadam-Hojjati, S., Barata, J.: A systematic review on multimodal emotion recognition: Building blocks, current state, applications, and challenges. *IEEE Access* **12**, 103976–104019 (2024). <https://doi.org/10.1109/ACCESS.2024.3430850>
9. Livingstone, S.R., Russo, F.A.: The ryerson audio-visual database of emotional speech and song (ravdess): A dynamic, multimodal set of facial and vocal expressions in north american english. *PLOS ONE* **13**(5), 1–35 (05 2018). <https://doi.org/10.1371/journal.pone.0196391>
10. Mollahosseini, A., Hasani, B., Mahoor, M.H.: Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing* **10**(1), 18–31 (2019). <https://doi.org/10.1109/TAFFC.2017.2740923>

11. Mollahosseini, A., Hassani, B., Salvador, M.J., Abdollahi, H., Chan, D., Mahoor, M.H.: Facial expression recognition from world wild web. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 1509–1516 (2016). <https://doi.org/10.1109/CVPRW.2016.188>
12. Ramirez Duque, A.A., Lindsay, A., Foster, M.E., Petrick, R.P.A.: A lightweight artificial cognition model for socio-affective human-robot interaction. In: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction. p. 929–933. HRI '24, Association for Computing Machinery, New York, NY, USA (2024). <https://doi.org/10.1145/3610977.3637469>
13. Rossi, A., Da Pozzo, E., Menicagli, D., Tremolanti, C., Priami, C., Sîrbu, A., Clifton, D.A., Martini, C., Morelli, D.: A public dataset of 24-h multi-levels psycho-physiological responses in young healthy adults. *Data* **5**(4) (2020). <https://doi.org/10.3390/data5040091>
14. Russell, J.A.: A circumplex model of affect. *Journal of Personality and Social Psychology* **39**(6), 1161–1178 (Dec 1980). <https://doi.org/10.1037/h0077714>
15. Samal, P., Hashmi, M.F.: Role of machine learning and deep learning techniques in eeg-based bci emotion recognition system: a review. *Artificial Intelligence Review* **57** (3 2024). <https://doi.org/10.1007/s10462-023-10690-2>
16. Siddiqui, M.F.H., Dhakal, P., Yang, X., Javaid, A.Y.: A survey on databases for multimodal emotion recognition and an introduction to the viri (visible and infrared image) database. *Multimodal Technologies and Interaction* **6** (6 2022). <https://doi.org/10.3390/mti6060047>
17. Subramanian, R., Wache, J., Abadi, M.K., Vieriu, R.L., Winkler, S., Sebe, N.: Ascertain: Emotion and personality recognition using commercial sensors. *IEEE Transactions on Affective Computing* **9**(2), 147–160 (2018). <https://doi.org/10.1109/TAFFC.2016.2625250>
18. Valentim, P., Muchaluat-Saade, D.: Affective tv: Concepts of affective computing applied to digital television. In: Marcus, A., Rosenzweig, E., Soares, M.M. (eds.) *Design, User Experience, and Usability*. pp. 203–220. Springer Nature Switzerland, Cham (2024)
19. Yoon, J., Kang, C., Kim, S., Han, J.: D-vlog: Multimodal vlog dataset for depression detection. *Proceedings of the AAAI Conference on Artificial Intelligence* **36**(11), 12226–12234 (Jun 2022). <https://doi.org/10.1609/aaai.v36i11.21483>
20. Younis, E.M.G., Zaki, S.M., Kanjo, E., Houssein, E.H.: Evaluating ensemble learning methods for multi-modal emotion recognition using sensor data fusion. *Sensors* **22**(15) (2022). <https://doi.org/10.3390/s22155611>
21. Yu, S., Androsov, A., Yan, H., Chen, Y.: Bridging computer and education sciences: A systematic review of automated emotion recognition in online learning environments. *Computers and Education* **220** (10 2024). <https://doi.org/10.1016/j.compedu.2024.105111>
22. Zheng, W.L., Lu, B.L.: Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development* **7**(3), 162–175 (2015). <https://doi.org/10.1109/TAMD.2015.2431497>

EEG e UX: Como a Análise das Ondas Cerebrais Pode Contribuir para a Área de Experiência do Usuário

Wesley Alves da Silva , Maristela de Freitas Riquelme ,
Giulia Pintan de Souza Sant'Ana , Arthur Henrique da Silva ,
Ana Vitória Maciel de Lima Pessoa  e Valdecir Becker 

Laboratório de Interação e Mídia, Centro de Informática, Universidade Federal da
Paraíba, João Pessoa, Paraíba, Brazil
contato@lim.ci.ufpb.br

Abstract. Aiming to overcome the limitations of subjective methods in User Experience (UX), which may introduce biases regarding participant perception, this research explored the use of neuroscience to measure viewer preference. Using electroencephalography (EEG), the brain electrical activity of 20 participants was monitored to analyze how live-action and animation advertisements modulate emotional responses. The analysis of Alpha and Beta waves demonstrated that live-action by employing a narrative with actors, achieved superior acceptance, inducing positive emotions and focus. In contrast, animation proved to be suitable for informational purposes. The study validates EEG as a robust tool for objective evaluation, providing metrics less susceptible to cognitive biases and demonstrating the efficacy of integrating neurophysiological data into UX practice.

Keywords: User Experience · Electroencephalography (EEG) · Neuroscience · Media Format · Emotional Response .

1 Introdução

O presente artigo investiga como o estudo das ondas neurais, por meio da eletroencefalografia (EEG), pode contribuir para a área de *User Experience* (UX), campo de estudo dedicado a avaliar a qualidade da interação entre indivíduos e sistemas, produtos ou serviços, de modo a garantir que estejam alinhados com as expectativas do usuário, sendo eficazes e satisfatórios. Nesse contexto, o objetivo da pesquisa é identificar padrões neurais de respostas emocionais para inferir preferências e receptividade dos participantes, propondo um método inovador e complementar às técnicas tradicionais utilizadas em UX.

Foram comparados dois vídeos promocionais de uma mesma loja com abordagens distintas: 1) um em *live-action* [1] (*storytelling* emocional) e 2) um institucional [2] (informativo e objetivo). Intentou-se verificar se essas abordagens causam efeitos emocionais diferentes e elucidar os fatores que garantem uma melhor experiência. Situada na Interação Humano-Computador (IHC), a pesquisa

integra UX, Interfaces Cérebro-Máquina (BCI) e neurociência. O experimento contou com 20 voluntários e utilizou uma metodologia mista (EEG e formulários pré e pós-teste) para avaliar métricas de felicidade, satisfação e aceitação.

A pesquisa foi metodologicamente construída com base na *Design Science Research* (DSR), estruturada nas seguintes etapas: (1) Identificação do problema e motivação, (2) Definição dos objetivos da solução, (3) Projeto e desenvolvimento, (4) Demonstração, (5) Avaliação e (6) Divulgação dos resultados e trabalhos futuros.

Sob a ótica da Experiência do Usuário, os resultados revelaram uma preferência significativa pelo vídeo *live-action*, confirmada pelos dados de EEG que apontaram menor carga cognitiva e maior relaxamento. Em contrapartida, a animação demandou maior esforço de processamento e ativação mental. Essa disparidade valida a eficácia do apelo emocional e demonstra que o cruzamento de dados neurofisiológicos (objetivos) e autorrelatados (subjetivos) constitui uma metodologia robusta para a avaliação da experiência do usuário frente a conteúdo midiático e, por conseguinte, a qualquer aplicação que envolva a análise de interações do usuário.

2 Atividade 1: Identificação do problema e motivação

A pesquisa em UX fundamenta-se em aplicações centradas no usuário e tradicionalmente emprega métodos subjetivos, qualitativos e quantitativos [3], como formulários, entrevistas, pesquisas online, testes de usabilidade e análise comportamental [4], para compreender a interação e avaliar a satisfação. Apesar de amplamente aceitas, essas técnicas apresentam limitações associadas à subjetividade, como vieses, imprecisões e influência de fatores externos, motivando a incorporação de abordagens tecnológicas capazes de acessar respostas cognitivas e emocionais de forma mais direta e objetiva.

A literatura recente evidencia o crescimento do uso de EEG como ferramenta não invasiva para capturar respostas emocionais e cognitivas em tempo real, superando limitações dos métodos autorrelatados [5]. No âmbito do estado da arte, a estimulação emocional por meio de vídeos é uma das abordagens em estudos de UX com EEG, validando o uso desse tipo de estímulo em contextos experimentais [5].

Diferentes estudos têm aplicado o EEG na avaliação de anúncios publicitários para prever valência e despertar emocional, demonstrando que os sinais neurais são mais eficazes para prever o comportamento do consumidor e caracterizar formatos que trazem mais conforto e retorno emocional frente ao conteúdo e a marca associado [6][7]. A comparação entre diferentes formatos de mídia também tem sido explorada na vanguarda da pesquisa em interação humano-computador. Ao investigar a carga cognitiva em interfaces adaptativas utilizando métricas como as bandas Alfa e Beta, similares às analisadas neste trabalho, observou-se que estruturas atípicas aumentam o esforço mental dos usuários [8]. Esse achado dialoga com a conclusão de que a animação institucional, focada em produtos, demandou maior processamento cognitivo do que o vídeo em *live-action*.

Outras pesquisas reforçam a relevância do uso de EEG de baixo custo para avaliação da experiência do usuário, consolidando-se como uma tendência na área. Em estudo que empregou o dispositivo *OpenBCI* para analisar a experiência de jogadores em diferentes condições sonoras, foram monitoradas as bandas Alfa e Beta para identificar estados de relaxamento e estresse, concluindo-se que a presença de som aumentou o estresse dos participantes [9].

Nesse contexto, o presente estudo estrutura sua metodologia às etapas típicas consolidadas na literatura: aquisição, pré-processamento, extração de características e análise, contribuindo com uma abordagem própria de visualização e interpretação dos dados.

Visando mitigar as lacunas originadas da subjetividade, o estudo analisou dois vídeos com propostas distintas: um *live-action* de 60 segundos, criado para o 30º aniversário do Lojão Rio do Peixe, que busca evocar pertencimento e cuidado por meio de cenas familiares e tom melodramático; e uma animação publicitária de 30 segundos, focada na exposição direta de produtos, preços e formas de pagamento, sem construção narrativa ou apelo emocional. A proposta foi comparar o apelo emocional e o engajamento gerados por esses formatos utilizando uma metodologia neurocientífica, atuando como complemento às limitações dos métodos tradicionais de UX [10]. A análise das ondas cerebrais possibilitou a obtenção de dados imediatos e objetivos sobre respostas emocionais e inconscientes, permitindo inferir com maior precisão o nível de satisfação associado à experiência do espectador [11].

3 Atividade 2: Definição dos objetivos da solução

A avaliação de interfaces e conteúdos midiáticos está associada à qualidade do produto e à satisfação do usuário, recorrendo aos próprios usuários como avaliadores [12]. Entretanto, métodos tradicionais baseados em percepções subjetivas podem sofrer influências de vieses, omissões e imprecisões [13], o que evidencia a necessidade de abordagens complementares para avaliar a experiência do usuário além do autorrelato.

Para atender a essa demanda, o estudo utiliza a eletroencefalografia, um método não invasivo que posiciona eletrodos em locais estratégicos do couro cabeludo para registrar variações na atividade cerebral decorrentes das trocas elétricas entre neurônios [14]. A pesquisa contou com 20 voluntários, que assistiram aos vídeos enquanto sua atividade cerebral era monitorada pelo equipamento Emotiv Insight [15].

A visualização e análise das ondas cerebrais foram realizadas por meio de um sistema [16] do Laboratório de Interação e Mídia (LIM) [17], permitindo a análise interativa dos dados em formato CSV, a aplicação de filtros de tempo e a avaliação tanto dos gráficos individuais de EEG quanto da média dos participantes, possibilitando uma análise global do experimento proposto.

4 Atividade 3: Projeto e desenvolvimento

Nesta fase, desenvolveu-se e aprimorou-se a metodologia de captação e classificação das flutuações neurais para inferência emocional. A etapa inicial consiste no recrutamento dos participantes, sem definição de público-alvo, já que o objetivo é identificar emoções percebidas independentemente de variáveis sociodemográficas, e na aplicação de um pré-teste sobre sono, fome e fadiga, os quais são considerados variáveis de confusão [18].

A medição das ondas neurais é realizada em ambiente controlado, com redução de fatores como ruídos, presença de terceiros e vieses sociais [19]. Na colocação da touca, observam-se morfologia capilar e circunferência craniana, pois cabelos volumosos ou perímetros pequenos ou muito expressivos podem comprometer a conectividade mínima de 80%. Após o acoplamento, aguarda-se de 30 a 40 segundos para a linha de base neural [20], seguido da apresentação dos estímulos com registro contínuo de EEG.

O estudo analisa ondas cerebrais associadas a estados mentais e emocionais, com foco nas bandas alfa (relaxamento) [21], beta baixa (concentração) e beta alta (excitação emocional intensa) [22] [23]. Ao final, um pós-teste coleta percepções subjetivas, como satisfação, felicidade, agitação, simpatia e intenção de recomendação, possibilitando correlacionar esses relatos às medidas fisiológicas e visualizar variações de engajamento, emoções e carga cognitiva, integrando métricas objetivas às técnicas tradicionais de UX.

5 Atividade 4 e 5: Demonstração e Avaliação

O teste foi realizado individualmente, em sala controlada. No pré-teste, coletaram-se dados demográficos, fatores que afetam a atividade cerebral (sono e uso de medicamentos) e escala Likert (1–5) para humor e condição física [24], considerando a influência desses estados na amplitude e consistência dos sinais de EEG [25]. Participaram 20 indivíduos (19–49 anos, predominância de 23 anos), com baixos níveis de cansaço e sonolência, dos quais dois faziam uso de medicamentos controlados.

A coleta neurofisiológica utilizou o Emotiv Insight 2 [15], seguindo o protocolo 10–20. A calibração via Emotiv Launcher foi considerada adequada com quatro dos cinco eletrodos acima de 80% de qualidade. Após estabilização e linha de base de 30–40 s, exibiram-se dois comerciais (*live-action* e animado) em sequência fixa, seguidos de um pós-teste com perguntas de preferência e escala Likert.

No pós-teste, 18 preferiram o *live-action* e 2 o animado. As médias foram: *live-action* — satisfação 3.5, felicidade 4.36, agitação 1.64, simpatia 4.29, recomendação 3.86; animação — satisfação 2.5, felicidade 2.79, agitação 2.64, simpatia 3.07, recomendação 2.36.

A análise de EEG considerou as faixas de onda alfa e beta. A alfa (8–12 Hz) relaciona-se a relaxamento e imersão [26] [27]; reduções indicam agitação [28]. A beta (13–30 Hz) corresponde a atenção e processamento: beta baixa (13–20 Hz) sinaliza foco equilibrado [29] beta alta (21–30 Hz) associa-se a excitação intensa e estresse. Valores muito baixos podem indicar desinteresse [30].

A Figura 1 ilustra o comportamento dos sinais neurais durante a exibição do vídeo 1, permitindo a observação das oscilações em tempo real.

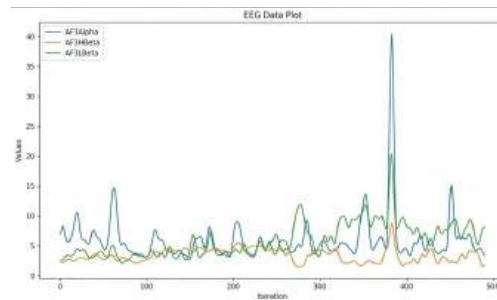


Fig. 1. Média das ondas captadas por EEG. [Gerada no Site do LIM]

No vídeo 1 (*live-action*), observou-se variação expressiva nos padrões de EEG. A faixa alfa (AF3Alfa) apresentou diversos picos ao longo dos 60 segundos, com destaque próximo a 48,75 s, indicando envolvimento emocional e introspectivo [30]. A beta baixa (AF3LBeta) manteve nível moderado e oscilou de forma sincronizada à alfa, sugerindo atenção contínua e engajamento cognitivo sem sobrecarga [31] [32]. A beta alta (AF3HBeta) permaneceu baixa, com pequenos picos apenas em momentos de maior excitação narrativa, indicando ausência de ansiedade ou esforço cognitivo excessivo [21]. Esses dados evidenciam experiência emocional intensa, foco constante e leve excitação, sem sinais de estresse.

Em contrapartida, a Figura 2 apresenta os dados coletados durante o segundo estímulo, evidenciando um padrão de resposta distinto.

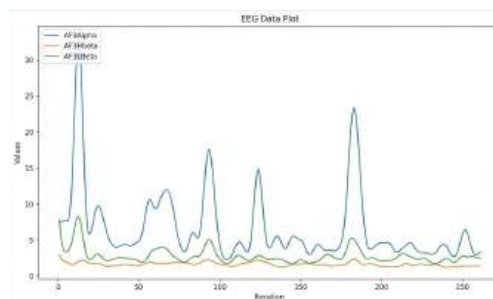


Fig. 2. Média das ondas captadas por EEG. [Gerada no Site do LIM]

No vídeo 2 (animação), o padrão EEG foi distinto. A alfa apresentou um pico inicial ($\approx 2,5$ s), seguido de níveis baixos e estáveis, indicando reação inicial de curiosidade e posterior relaxamento passivo [27]. As faixas beta baixa e beta alta permaneceram lineares, revelando atenção superficial e ausência de processamento cognitivo profundo ou ativação emocional complexa [32]. 

A comparação entre os gráficos evidencia diferenças claras entre as propostas comunicacionais. O vídeo 1 gerou um padrão EEG mais dinâmico, emocional e cognitivamente ativo — com picos de alfa ligados à imersão, beta baixa oscilante indicando foco e ausência de beta alta — enquanto o vídeo 2 mostrou resposta mais estável, marcada por atenção passiva e baixa ressonância emocional. Esses achados apontam maior eficácia emocional e comunicacional do vídeo 1 e reforçam que dados neurofisiológicos mensuram a experiência do usuário de forma mais precisa. As discrepâncias revelam que a eficácia em UX depende da capacidade do estímulo de gerar empatia, sendo o *live-action* favorecido pela ativação de mecanismos de identificação humana ausentes na animação.

Nesse contexto, a relação com a Experiência do Usuário torna-se evidente: o conteúdo audiovisual atua como interface de engajamento, e estímulos afetivamente ricos reduzem a carga de processamento e ampliam a retenção. Dados neurofisiológicos, portanto, permitem compreender com maior precisão como o usuário responde aos diferentes tipos de estímulo.

6 Atividade 6: Divulgação dos resultados e trabalhos futuros

O presente estudo utilizou o EEG para captar a atividade elétrica cerebral e avaliar objetivamente as respostas emocionais de indivíduos expostos a vídeos em *live-action* e animação. O objetivo foi estabelecer uma ponte rigorosa com a área de UX, oferecendo um meio inovador para complementar os métodos tradicionais que, pela sua natureza subjetiva, apresentam limitações na mensuração do engajamento e da receptividade.

A pesquisa envolveu 20 participantes, utilizando questionários pré e pós-teste para validar as interpretações das ondas neurais registradas. Os resultados apontaram uma preferência evidente pelo formato em *live-action*, que demonstrou desempenho superior tanto nas métricas subjetivas (satisfação, simpatia e recomendação) quanto no engajamento neural mais duradouro. Essa capacidade de compreender as preferências de maneira objetiva viabiliza o desenvolvimento de produtos mais eficazes, fornecendo uma base robusta para a tomada de decisão de profissionais de design sem substituir as práticas de UX já consolidadas.

Contudo, foram identificados desafios técnicos significativos durante a coleta. O equipamento de EEG demonstrou pouca flexibilidade em circunferências cranianas menores e desempenho reduzido em indivíduos com cabelos volumosos ou espessos, dificultando o contato dos eletrodos com o couro cabeludo. Apesar dessas restrições, o rigor metodológico empregado assegurou captações satisfatórias para uma interpretação confiável das ondas cerebrais.

Para a continuidade da pesquisa, considera-se indispensável ampliar a diversidade da amostra e realizar experimentos com estímulos interativos, como aplicativos e websites. Recomenda-se, ainda, investigar a relação entre o estado inicial dos usuários (pré-teste) e sua resposta emocional durante a experiência. Em última análise, este estudo validou sua originalidade sob a abordagem da DSR, consolidando a integração entre EEG e UX como um método eficaz para a mensuração objetiva da preferência do usuário em relação a produtos digitais.

Referências

1. Lojão Rio do Peixe: Semana do Cliente Lojão Rio do Peixe 2024. YouTube, <https://www.youtube.com/watch?v=4ZpTEtDQ1tA>. Acessado em 10/12/2025
2. Lojão Rio do Peixe: Filme institucional 2024 (LRP). YouTube, https://www.youtube.com/watch?v=K_noCayM-9w. Acessado em 26 jan 2025
3. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
4. Tullis, T., Albert, W.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. 2nd edn. Morgan Kaufmann, Burlington (2008)
5. Leite Filho, J.R., Coleti, T.A., Morandini, M.: Evaluating chatbot user experience (UX) through electroencephalography measures: A systematic literature review. Computers in Human Behavior Reports **13**, 100381 (2024). <https://doi.org/10.1016/j.chbr.2024.100381>
6. Shukla, A., Gullapuram, S.S., Katti, H., Kankanhalli, M., Winkler, S., Subramanian, R.: Recognition of Advertisement Emotions With Application to Computational Advertising. IEEE Transactions on Affective Computing **13**(2), 781–792 (2022). <https://doi.org/10.1109/TAFFC.2020.2964549>
7. Lima, M., Pinon, A., Fontes, R., Alves, W., Macedo, J., Becker, V.: Analysis of the Enjoyment Levels Evoked by Two Video Commercials Using EEG. In: Kurosu, M., Hashizume, A. (eds) Human-Computer Interaction. HCII 2025. Lecture Notes in Computer Science, vol 15768. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-93845-0_15
8. Figueiredo, G.D., Abrahão, S., Insfran, E., Vanderdonckt, J.: Measuring User Experience of Adaptive User Interfaces using EEG: A Replication Study. arXiv preprint (2023). <https://doi.org/10.48550/arXiv.2306.03525>
9. Cano, S., Araujo, N., Guzman, C., Rusu, C., Albiol-Pérez, S.: Low-Cost Assessment of User eXperience Through EEG Signals. IEEE Access **8**, 158475–158487 (2020). <https://doi.org/10.1109/ACCESS.2020.3017685>
10. Goodman, E., Kuniavsky, M., Moed, A.: Observing the User Experience: A Practitioner's Guide to User Research. 2nd edn. Morgan Kaufmann (2012)
11. Şahin, E.B., Durdu, P.O.: Evaluating the potential of using EEG-based BCI method in user experience research. Universal Access in the Information Society **24**(3), 2507–2530 (2025). <https://doi.org/10.1007/s10209-025-01207-5>
12. ISO: ISO 9241-210: Ergonomia da interação humano-sistema — Parte 210: Design centrado no ser humano para sistemas interativos. 2ª ed. ISO (2019)
13. Nisbett, R.E., Wilson, T.D.: Telling more than we can know: Verbal reports on mental processes. Psychological Review **84**(3), 231–259 (1977). <https://doi.org/10.1037/0033-295X.84.3.231>
14. iMotions: Unpack Human Behavior: Electroencephalography – The Complete Pocket Guide (2019). <https://imotions.com/blog/learning/best-practice/eeg/>, último acesso em 02/12/2025

15. EMOTIV: Insight – 5 Channel Wireless EEG Headset. <https://www.emotiv.com/insight/>. Acessado em 09/12/2025
16. Becker, V., Cavalcanti, M., Silva, T., et al.: A System for Graphical Visualization of Brainwaves to Analyse Media Content Consumption. In: Lecture Notes in Computer Science. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05409-9_24
17. Laboratório de Interação e Mídia. <https://lim.tec.br>, último acesso em 22/09/2025
18. Shadish, W.R., Cook, T.D., Campbell, D.T.: Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Mifflin, Boston (2002)
19. Luck, S.J.: An Introduction to the Event-Related Potential Technique. 2nd edn. MIT Press, Cambridge (2014)
20. Cacioppo, J.T., Tassinary, L.G., Berntson, G.G.: Handbook of Psychophysiology. 3rd edn. Cambridge University Press, Cambridge (2007)
21. Lopes da Silva, F.H.: EEG and MEG: basis, principles, and applications. In: Handbook of Clinical Neurology, pp. 129–165. Elsevier (2013)
22. Pfurtscheller, G., Lopes da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology* **110**(11), 1842–1857 (1999)
23. Başar, E., Başar-Eroglu, C., Karakaş, S., Schürmann, M.: Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology* **39**(2-3), 241–248 (2001)
24. Likert, R.: A technique for the measurement of attitudes. *Archives of Psychology* **22**(140), 1–55 (1932)
25. Steriade, M., McCarley, R.W.: Brain Stem Control of Wakefulness and Sleep. Plenum Press, New York (1990)
26. Niedermeyer, E.: The normal EEG of the waking adult. In: *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Williams & Wilkins, Baltimore (1999)
27. Klimesch, W.: Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences* **16**(12), 606–617 (2012)
28. Palva, S., Palva, J.M.: New vistas for alpha-frequency band oscillations. *Trends in Neurosciences* **30**(4), 150–158 (2007)
29. Klimesch, W.: EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews* **29**(2-3), 169–195 (1999)
30. Basar, E.: A review of alpha activity in integrative brain function. **86**(1), 1–24 (2012)
31. Knyazev, G.G.: Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews* **31**(3), 377–395 (2007)
32. Harmony, T.: The functional significance of delta, theta, alpha, beta and gamma oscillations in the brain. *Arquivos de Neuro-Psiquiatria* **71**(9B), 621–625 (2013)

Interactive Video Experiences



Interactive Shoppable Video Strategies for Specific Markets: A Focus on Portugal and Spain

Sudarshan Kadam¹ [0009-0003-4400-787X], David Badajoz² [0000-0003-3428-3312], Telmo Silva¹ [0000-0001-9383-7659], and Jorge Abreu¹ [0000-0002-0492-2307]

¹ Digimedia, Department of Communication and Art, University of Aveiro, Portugal
sudarshan@ua.pt, tsilva@live.ua.pt, jfa@ua.pt

² Universitat Autònoma de Barcelona, Spain
David.Badajoz@uab.cat

Abstract. Connected devices and streaming platforms are reshaping media consumption by enabling interactive, shoppable videos that facilitate real-time product exploration and instant buying. This is an interactive shoppable format that merges entertainment and commerce, allowing viewers to engage directly and purchase featured products within the viewing experience. It has the potential to enhance engagement and purchase intent by streamlining the shopping journey through real-time interactivity and context discovery, leading to purchases via an embedded link, with its effectiveness varying across markets due to behavioral, technological, and cultural factors. In this framework, this study examines Portugal and Spain to analyze adoption, user acceptance, and trust in embedded purchasing features. Emerging evidence points to the potential benefits of e-commerce, such as greater personalization and ease of use. Also underscoring challenges related to production complexity, privacy concerns, and the risks of deceptive practices, this paper focuses on Portugal and Spain, introducing three conceptual contributions: (1) Iberian Shoppable Video Acceptance Model (ISVAM), (2) Iberian Interactive Commerce Personas, and (3) Portugal–Spain Comparative Matrix designed to capture behavioral, cultural and technological determinants of interactive commerce adoption. ISVAM is proposed as a conceptual framework, with empirical validation planned through survey-based modelling and structural equation modelling (SEM) in future research. Together, these contributions will establish a framework for analyzing the adoption, trust, and experience of interactive video interfaces among Iberian audiences.

Keywords: Interactive video, Interactive ads, Shoppable video, Iberian markets, Embedded purchase, Connected TV ecosystem, Portugal, Spain, User Engagement, Purchase behavior

1 Introduction

The evolution of connected devices, smart televisions, and streaming ecosystems has reshaped how audiences interact with digital media [1]. Interactive features such as clickable links (e.g., tapping a product tag to open item details), dynamic overlays (e.g., product cards appear at specific intervals in the video), platform-driven recommendations (e.g., showing similar items based on viewing history) and embedded checkout pathways (e.g., allowing users to add products without leaving the video player) have transformed passive viewing into transactional engagement. Despite the evolution of interactive video technologies, a limited number of studies examine how cultural differences, privacy expectations, and platform maturity shape adoption in Iberian markets, revealing a clear need for a region-specific acceptance framework. Research demonstrates that these features can stimulate impulse buying in short-video environments by influencing emotional arousal, perceived control, and purchase intention [2] [3] [4] [5]. This aligns with core predictions from the Technology-Adoption Model (TAM; usefulness and ease of use) and the Unified Theory of Acceptance and Use of Technology (UTAUT; performance and effort expectancy) [6] [7] [8].

A recent review of interactive videos identifies interaction, trust, platform design, and social influence as the dominant behavioural roles of usability and social influence in technology [9] [1] [10]. Portugal and Spain represent an effective comparative market pair. Although they are geographically adjacent, they exhibit notable differences in privacy sensitivity and digital risk perception [11], influencer-driven behaviour, platform adoption, and purchase readiness. These distinctions are consistent with regional studies demonstrating variation in channel choice, trust formation, and social-scene characteristics across digital commerce environments [11] [12].

This study adopts a conceptual, exploratory approach grounded in a literature synthesis and preliminary observations that inform the development of personas and comparative insights. Empirical validation is planned in subsequent phases through survey-based data collection and experimental testing.

2 Literature review

Interactive shoppable videos draw from four main domains: interactive media, HCI, digital commerce, and consumer behaviour. These domains form the conceptual foundations for the subsequent review, which examines how interactivity, commercial formats, trust mechanisms, and regional dynamics influence viewer behaviour. In commerce-oriented video environments, interactive pathways such as product overlays, taps, or swipes strengthen emotional and behavioral engagement [2] [13]. These effects are well documented in short-video shopping contexts, where interactivity directly predicts purchase intention [3].

Shoppable media emerged from early hypervideo forms and evolved through tagging, dynamic overlays, and recommendation algorithms. Embedded purchasing reduces resistance by integrating discovery and checkout into a single flow. [2] [4]. However, challenges persist, including interface overload, distraction, and tensions between persuasive design and entertainment value [11]. In shoppable videos, trust is shaped by

platform credibility but also by transparency surrounding data practices, recommendation logic, and privacy safeguards [12]. GDPR strengthens user rights but increases expectations for clarity, informed consent, and non-manipulative design.

Iberian Adoption Patterns show that Portuguese consumers exhibit heightened privacy concerns, while Spanish audiences respond more strongly to influencer-driven and entertainment-centric content [14] [5]. To contextualize behavioural differences between Portugal and Spain, Hofstede's Cultural dimensions offer a useful macro-level lens for explaining variations in privacy sensitivity, creator influence, and openness to commercial formats [15] [16].

The review suggests that conventional technology-adoption frameworks may inadequately reflect the culturally specific and privacy-sensitive aspects of shoppable video usage in Iberia, highlighting the necessity for a specialized conceptual framework, the Iberian Shoppable Video Acceptance Model (ISVAM).

Recent work on OTT interfaces and interactive television highlights the importance of usability constraints, remote-control interaction, and cognitive load in shaping user behaviour, suggesting that interactive commerce must be designed within device-specific limitations

3 Proposed Model and Theoretical Framework

ISVAM is proposed as a **conceptual framework at this stage**, with empirical validation planned through survey-based modelling and Structural Equation Modelling (SEM) in future work.

ISVAM is proposed as a conceptual framework that integrates behavioural, technological, and cultural factors that influence the adoption of interactive shopping video formats in Portugal and Spain. The Iberian Shoppable Video Acceptance Model (ISVAM) builds upon established technology adoption frameworks while addressing their limitations in interactive commerce contexts. The Technology Acceptance Model (TAM) explains adoption through perceived usefulness and ease of use, while UTAUT extends this by incorporating social influence and facilitating conditions within platform ecosystems. However, these models do not fully capture the complexities of interactive shoppable video environments. ISVAM extends these frameworks by integrating interactivity, trust, privacy sensitivity, and cultural context, making it more suitable for analyzing adoption in region-specific settings such as Portugal and Spain.

While grounded in established technology-adoption theories, including the Technology Acceptance Model (TAM) [6] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [8]. ISVAM extends these models by incorporating privacy expectations, cultural dynamics, and platform-specific interaction patterns relevant to Iberian digital markets.

Unlike TAM and UTAUT, which primarily focus on functional and social determinants of technology adoption, ISVAM explicitly integrates cultural alignment and privacy assurance as context-sensitive moderators, addressing gaps in explaining user behaviour in regulated and culturally diverse markets such as Iberia.

ISVAM Construct

- Perceived Interactivity (PI): PI is the user's sensed level of control, two-way communication, and responsiveness in the shoppable video experience [2].
- Perceived Usefulness (PU): PU is the degree to which shoppable video features improve user convenience, efficiency, and decision-making [2] [14].
- Trust & Privacy Assurance (TPA): TPA is the user's confidence in the platform's data protection, transparency, and overall credibility [17].
- These hypothesised paths will be empirically examined using Structural Equation Modelling (SEM), including multi-group analysis [16] for Portugal and Spain.

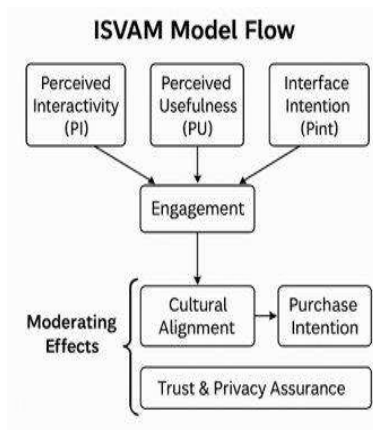


Fig. 1. ISVAM Flow

Building on the literature review, the ISVAM framework adapts and integrates core constructs from TAM and its successors. Perceived usefulness maps directly onto ISVAM's Perceived Usefulness (PU), while perceived ease of use aligns with Interface Usability (IU). UTAUT's expanded constructs, including performance expectancy, effort expectancy, and social influence, are reflected in perceived interactivity, Cultural Alignment, and engagement within ISVAM. Privacy assurance and cultural factors are highly relevant in markets such as Portugal and Spain.

Hypothesis:

The hypothesis is based on previous research conducted in TAM [6], UTAUT [17], and technology-adoption extensions [7]. This study develops a set of hypotheses to examine how interactivity, usefulness, usability, trust, privacy assurance, and cultural fit influence user engagement and purchase intention in shoppable video. The proposed ISVAM-UTAUT integrated framework assumes the following structural relationships:

- H1, H2, H3: Perceived Interactivity (PI), Perceived Usefulness (PU), and Interface Usability (IU) each positively influence Engagement.
- H4, H5: Engagement is a good predictor of purchase intention.

- H6, H7: Cultural Alignment (CA) and Trust & Privacy Assurance (TPA) moderate the relationship between Engagement and Purchase Intention, strengthening this effect when CA or TPA is high.
- H8, H9, H10, H11: Platform-related conditions, represented through UTAUT constructs such as social influence, effort expectancy, performance expectancy, and facilitating conditions, are expected to influence PI, PU, and IU.

4 Iberian Interactive Commerce Personas

These personas are derived from a theoretical synthesis of prior literature and exploratory insights, and are intended as analytical archetypes rather than empirically validated user segments

Derived from the literature, four personas illustrate behavioral profiles in Portugal and Spain. These personas operationalize ISVAM by showing how variations in trust, cultural alignment, interactivity expectations, and usability preferences manifest.

Persona A, the Skeptical Explorer from Portugal, stands for privacy-conscious users who engage cautiously with shoppable video features. These individuals prefer non-intrusive experiences when privacy concerns are significant [11]. Trust signals are carefully assessed [12].

Persona B, the Experience-Driven Shopper from Spain, demonstrates a strong affinity for entertainment-rich, creator-led content. These consumers respond with high enthusiasm to storytelling and influencer-based product demos, and show high receptivity to narrative-driven persuasion [12], and have a wider entertainment orientation [14].

Persona C, the Efficiency Seeker, represents users in Portugal and Spain who prioritize functional value and convenience in their interactions. These individuals favour fast, frictionless experiences, are sensitive to cognitive loads and interface complexity [9], and are less influenced by branding elements or storytelling-based persuasive cues.

Persona D, also known as the Influence-Follower, primarily represents Spanish users who engage through social validation mechanisms. They rely heavily on community-based cues such as reviews, comments, and creator recommendations, making them particularly susceptible to social-signal-driven persuasion [2] [13], especially within TikTok and Instagram-native shoppable formats.

The personas align closely with national-level tendencies reflected in Hofstede's cultural dimensions [15], providing a macro-level explanation for differences in privacy sensitivity, creator influence, and openness to commercial formats.

5 Portugal–Spain theoretical Comparative Matrix

This comparative matrix is informed by behavioural insights and also by national-level tendencies identified through Hofstede's cultural dimensions, which explain contrasting levels of trust, individualism, uncertainty avoidance, and indulgence in Portugal and Spain [15] [16]. The comparative patterns outlined in Table 1 provide the empirical grounding for interpreting behavioral tendencies discussed below.

Table 1. Portugal-Spain comparative study matrix

Parameters	Portugal	Spain
Trust in digital ads	Medium	Medium-High
Privacy sensitivity	High	Moderate
Influence of creators	Medium	High
Streaming consumption	High	Very High
Conversion willingness	Low	High
Openness to commercial formats	Cautious	Experience-driven

Research indicates that trust, privacy concerns, and platform transparency weigh more heavily for Portuguese users[13]. In contrast, Spanish users are more substantially aligned with influencer-driven and entertainment-rich commerce formats[14].

6 Discussion

The review findings highlight several important considerations for Hofstede, particularly in uncertainty avoidance and indulgence, which help explain Portugal's cautious stance and Spain's stronger responsiveness to creator-driven formats[15] [16]. First, perceived interactivity and interface clarity strongly shape user engagement and purchase intention, consistent with evidence that interactive elements significantly influence consumer responses in digital commerce environments[2] [13] [3].

Second, technological readiness varies across platforms, particularly in smart TV ecosystems where support for interactive overlays remains inconsistent. Similar fragmentation has been identified in streaming and short-video commerce infrastructures, suggesting a need for platform-level harmonization.

Third, cultural differences significantly influence user motivation. Portuguese users prioritize transparency and data protection, aligning with studies linking trust, privacy concerns, and purchase intentions. Spanish users respond favorably to influencer-driven and entertainment-oriented content, echoing findings on influencer selection and social scene effects [14] [5]. These emerging tendencies reinforce ISVAM's theoretical foundations, particularly the roles of interactivity (PI), usability (IU), cultural fit (CA), and Trust and privacy assurance (TPA). These patterns support the structure of ISVAM, which positions Cultural Alignment and Trust & Privacy Assurance as key moderating constructs shaping engagement and purchase intention.

From an implementation perspective, interactive video systems face device-specific constraints, particularly in Smart TV environments where interaction relies on remote controls, limiting precision and increasing navigation effort. Additionally, overlay density and timing introduce cognitive load, potentially reducing usability and engagement. Platform fragmentation across OTT ecosystems further complicates the consistent implementation of interactive features.

To operationalise ISVAM, a typical system would include a content layer (video + metadata tagging), interaction layer (overlays, triggers), a recommendation engine, and transaction layer (embedded checkout).

This architecture highlights the technical feasibility and integration requirements of shoppable video systems.

These findings reinforce ISVAM's theoretical structure and align with TAM and UTAUT, while extending them through culturally embedded and privacy-sensitive constructs.

7 Future Work:

Future work will involve survey-based measurement development, cross-market data collection in Portugal and Spain, and validation using SEM and multigroup analysis, enabling refinement of ISVAM and comparison with TAM/UTAUT,

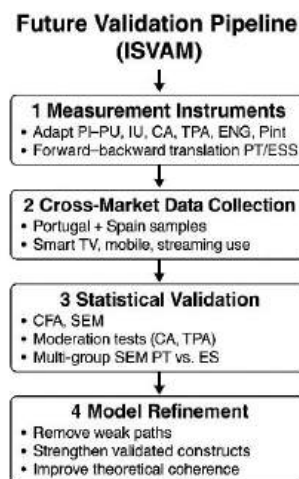


Fig. 2. Future work – ISVAM validation process

Measurement instruments for ISVAM constructs will be adapted from validated TAM [6], UTAUT [8], and TAM-extension scales [7]. Cross-market data will be collected from digital users in Portugal and Spain. For Statistical Validation, Model testing will use SEM for hypothesis testing and Multigroup SEM to compare Portugal and Spain. ISVAM will be iteratively refined by removing weak paths and strengthening relationships to ensure theoretical coherence and empirical robustness. A validated ISVAM model will be extended by integrating cultural alignment and GDPR-driven trust constructs, thereby improving understanding of Iberian shopping video behavior.

Final Remark

This paper lays the groundwork for understanding the adoption of shoppable video in Portugal and Spain, serving as a seed planted in the interactive shoppable video space. The proposed ISVAM model captures key behavioral, cultural, and technological factors shaping Iberian users' engagement, emphasizing the importance of interactivity, usability, trust, and cultural fit.

The next steps will involve empirical validation of ISVAM through survey-based modelling and comparative analysis across markets. This evolving work will culminate in a complete publication that will support academia, policymakers, and media organizations in creating user-centered, ethically responsible, and culturally attuned shoppable video experiences in Europe.

Acknowledgement and Funding

This work is supported by the University of Aveiro and the Universitat Autònoma de Barcelona (UAB) and was partially supported by the Erasmus+ MindPhD mobility grant and by additional institutional support from the University of Aveiro.

References

- [1] B. Cardoso, J. Abreu, and A. Prata, "The TV Ecosystem is Fragmented and in Need of New Business and Technological Strategies," in *ACM International Conference on Interactive Media Experiences*, Aveiro JB Portugal: ACM, June 2022, pp. 133–140. doi: 10.1145/3505284.3529976.
- [2] Y. Zhang, T. Zhang, and X. Yan, "Understanding impulse buying in short video live E-commerce: The perspective of consumer vulnerability and product type," *Journal of Retailing and Consumer Services*, vol. 79, p. 103853, July 2024, doi: 10.1016/j.jretconser.2024.103853.
- [3] Y. Guo* *et al.*, "Analyzing and Predicting Consumer Response to Short Videos in E-Commerce," *ACM Trans. Manage. Inf. Syst.*, vol. 15, no. 4, pp. 1–23, Dec. 2024, doi: 10.1145/3690393.
- [4] X. Luo, W. M. Lim, J.-H. Cheah, X.-J. Lim, and Y. K. Dwivedi, "Live Streaming Commerce: A Review and Research Agenda," *Journal of Computer Information Systems*, vol. 65, no. 3, pp. 376–399, May 2025, doi: 10.1080/08874417.2023.2290574.
- [5] C. Girsang and C.-H. Teng, "Exploring User Engagement and Purchase Intentions in T-Shirt Retail Through Augmented Reality and Instagram Filters," *Applied Sciences*, vol. 15, no. 18, p. 10161, Sept. 2025, doi: 10.3390/app151810161.
- [6] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, Sept. 1989, doi: 10.2307/249008.
- [7] N. Marangunić and A. Granić, "Technology acceptance model: a literature review from 1986 to 2013," *Univ Access Inf Soc*, vol. 14, no. 1, pp. 81–95, Mar. 2015, doi: 10.1007/s10209-014-0348-1.
- [8] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology1," *MIS Quarterly*, vol. 36, no. 1, pp. 157–178, Mar. 2012, doi: 10.2307/41410412.
- [9] C. Johnson, *Online TV*, 1st ed. London ; New York : Routledge, 2019.: Routledge, 2019. doi: 10.4324/9781315396828.
- [10] J. Abreu, P. Almeida, B. Teles, and M. Reis, "Viewer behaviors and practices in the (new) television environment," in *Proceedings of the 11th European Conference on Interactive TV and Video*, Como Italy: ACM, June 2013, pp. 5–12. doi: 10.1145/2465958.2465970.
- [11] F. Skwara, "Effects of mental accounting on purchase decision processes: A systematic review and research agenda," *J of Consumer Behaviour*, vol. 22, no. 5, pp. 1265–1281, Sept. 2023, doi: 10.1002/cb.2193.
- [12] W. Li, S. Cujilema, L. Hu, and G. Xie, "How Social Scene Characteristics Affect Customers' Purchase Intention: The Role of Trust and Privacy Concerns in Live Streaming Commerce," *JTAER*, vol. 20, no. 2, p. 85, Apr. 2025, doi: 10.3390/jtaer20020085.
- [13] X. Liu and L. Zhang, "Impacts of different interactive elements on consumers' purchase intention in live streaming e-commerce," *PLoS ONE*, vol. 19, no. 12, p. e0315731, Dec. 2024, doi: 10.1371/journal.pone.0315731.
- [14] F. Ye, L. Ji, Y. Ning, and Y. Li, "Influencer selection and strategic analysis for live streaming selling," *Journal of Retailing and Consumer Services*, vol. 77, p. 103673, Mar. 2024, doi: 10.1016/j.jretconser.2023.103673.

- [15] G. Hofstede, G. J. Hofstede, and M. Minkov, *Cultures and organizations: software of the mind: intercultural cooperation and its importance for survival*, Revised and Expanded third edition. New York: McGraw-Hill, 2010.
- [16] F. Sharmin, M. T. Sultan, D. Wang, A. Badulescu, and B. Li, “Cultural Dimensions and Social Media Empowerment in Digital Era: Travel-Related Continuance Usage Intention,” *Sustainability*, vol. 13, no. 19, p. 10820, Sept. 2021, doi: 10.3390/su131910820.
- [17] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User Acceptance of Information Technology: Toward A Unified View1,” *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, Sept. 2003, doi: 10.2307/30036540.



Pausa para pensar: Análisis de la Demanda Cognitiva y la Intencionalidad Didáctica en Videos Interactivos.

Miranda A.,^{1,2} Santos G.^{1,2}, Abásolo Guerrero M.J.^{2,3,4}

¹Educación en Ciencias con Tecnología (ECienTec), Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN), Argentina

²Departamento de Formación Docente, Facultad de Ciencias Exactas, UNICEN, Argentina
{nsantos,amiranda}@ecientec.exa.unicen.edu.ar

³Instituto de Investigación en Informática (III LIDI), Facultad de Informática, Universidad Nacional de la Plata (UNLP), Argentina

⁴Comisión de Investigaciones Científicas de la Pcia. de Bs. As. (CICPBA), Argentina
mjabasolo@lidi.info.unlp.edu.ar

Resumen. El trabajo presenta un estudio sobre las intervenciones pedagógicas que, en el marco de un curso de formación continua en educación mediada por tecnología, realizan docentes de matemática, física e informática para transformar videos tradicionales en videos educativos interactivos. Las preguntas insertas en un video interactivo brindan experiencia de aprendizaje activo y generan diferentes demandas cognitivas dependiendo del momento en que se ubican y del propósito pedagógico que las orienta. Se estudia cuáles son las estrategias de interacción que este grupo de docentes integra en videos disponibles en la Web mediante la utilización del software libre eXeLearning y la versión gratuita de Edpuzzle, a fin de identificar y caracterizar el tipo de interacción en relación a la intencionalidad didáctica y la demanda cognitiva que requiere. Se realiza un análisis cualitativo de las producciones de los cursantes, focalizando en las características de interacciones agregadas y las demandas cognitivas consideradas. Los resultados revelan un claro predominio de intervenciones basadas en preguntas de opción única y rellenar huecos. Las percepciones relevadas confirman el potencial pedagógico de los videos interactivos y muestran que la planificación de tareas interactivas, atendiendo a las demandas cognitivas, favorece la reflexión docente y amplía las estrategias para enriquecer los aprendizajes mediados por tecnologías.

Palabras clave: Videos interactivos, Demanda cognitiva, Aprendizaje activo, Formación docente.

1 Introducción

Los videos educativos interactivos son recursos que van más allá de la simple visualización, permiten al estudiante participar activamente mediante pequeñas tareas dentro del video. Se caracterizan por ofrecer un mensaje enriquecido a partir de la integración de textos, gráficos, audio, videos o diferentes desafíos en formato de preguntas. Tales desafíos o intervenciones requieren distintas demandas cognitivas e inciden en el procesamiento y la comprensión del contenido, dependiendo del momento en que se ubican y del propósito pedagógico que las orienta.

El diseño de videos educativos interactivos, ya sea a partir de videos breves disponibles o de producciones realizadas por el propio docente, requiere una cuidadosa planificación de las tareas o intervenciones que se incorporan con el propósito de favorecer un mayor involucramiento del estudiante. Esto motiva el desarrollo de nuevas estrategias de enseñanza tendientes al aprovechamiento del potencial de los videos interactivos que involucran componentes de aprendizaje activo en las tareas para promover el aprendizaje autónomo [1].

Los trayectos de formación docente continua en prácticas de enseñanza mediadas por tecnologías son instancias para actualizar el conocimiento sobre la potencialidad de las tecnologías, cuyos avances son permanentes y que impulsan la innovación en metodologías de enseñanza. Una variable importante a considerar en el diseño de videos interactivos es la sobrecarga de la memoria de trabajo del estudiante [2], para lo que es necesario evaluar la demanda cognitiva [3] que generan las tareas de aprendizaje.

2 Perspectiva Conceptual

En el marco de las metodologías activas de enseñanza y aprendizaje [4], se espera que las tareas enriquecidas con desafíos cognitivos “interactivos” favorezcan la motivación y faciliten la comprensión de los contenidos que se enseñan. La tecnología se constituye en un aliado estratégico para implementar este tipo de metodologías, no solo por su componente motivador, sino también por su potencial para presentar información de manera dinámica e involucrar al estudiantado en procesos cognitivos como interpretar, relacionar, reflexionar, explicar, sintetizar y concluir. En particular, los videos interactivos son valiosos recursos para desarrollar estrategias de enseñanza que involucren componentes de aprendizaje activo en las tareas [1].

Los principios de la teoría cognitiva de aprendizaje multimedia [5] son un marco adecuado para pensar y fundamentar lineamientos de diseño de videos interactivos [2]. Esta teoría supone que las personas procesan la información verbal y visual por separado; poseen capacidad limitada de procesamiento, por lo que tienen que seleccionar el material a procesar en la memoria de trabajo; organizan estructuras verbales y visuales e integran entre sí y con el conocimiento que consideran relevante de la memoria a largo plazo [3].

Cada tarea de aprendizaje demanda una peculiar carga cognitiva, asociada al nivel de pensamiento y a las operaciones mentales requeridas. Estas demandas pueden clasificarse de baja complejidad, cuando involucran procedimientos conocidos y habilidades básicas como atención, percepción, memoria o aplicación rutinaria de métodos, o de alta complejidad cuando exigen procesos de pensamiento avanzado como comprensión, análisis, la resolución de problemas, la toma de decisiones o la creatividad [6]. Manrique [7] propone una herramienta de análisis de situaciones de enseñanza que permite identificar los procesos que los estudiantes ponen en juego al realizar actividades de aprendizaje, así como describir dichas actividades en términos de las demandas cognitivas que le requiere a los participantes.

El diseño e implementación de propuestas de enseñanza mediadas por tecnología que promueven diversas formas de pensamiento es posible mediante la articulación integrada de conocimientos tecnológicos, pedagógicos y disciplinares. Este enfoque, sistematizado en el modelo TPACK (Technological Pedagogical Content knowledge) [8], propone que el uso educativo de la tecnología no depende solo del dominio técnico del docente, sino su integración significativa con los propósitos pedagógicos y los contenidos disciplinares. En otras palabras, no se trata de agregar conocimientos tecnológicos al saber pedagógico y disciplinar, sino de generar un tipo de conocimiento específico que solo se configura en la interrelación estrecha entre estos tres componentes.

El diseño de videos educativos interactivos constituye un ejemplo claro de una tarea de aprendizaje activo que pone en evidencia el conocimiento TPACK del docente. En la elaboración de estos videos, el docente debe crear intervenciones pedagógicas que generen distintas demandas cognitivas, adecuadas al momento del video en que se insertan y al propósito didáctico que las orienta.

El presente estudio analiza las estrategias de interacción que un grupo de docentes, en el marco de un curso de formación continua sobre mediación tecnológica, integra en videos de la Web. El objetivo es identificar y caracterizar dichas interacciones en relación con su intencionalidad didáctica y la demanda cognitiva que promueven.

3 Metodología

Este estudio de carácter cualitativo indaga en las estrategias pedagógicas que los docentes de la cohorte 2025 de la Diplomatura Universitaria Superior en Enseñanza Mediada por Tecnologías (DUSEMT) despliegan al diseñar videos interactivos.

El objetivo de la DUSEMT es formar a los docentes para planificar actividades de enseñanza con mediación tecnológica; comprender las transformaciones que deben operarse sobre los contenidos disciplinares y reflexionar sobre la propia práctica de enseñanza. La certificación se alcanza con un trabajo final integrador que consiste en el diseño colaborativo de una propuesta de enseñanza y la creación de un material didáctico digital (MDD) que incluye videos interactivos y la justificación pedagógica de las intervenciones agregadas [1].

El MDD se realiza con la aplicación eXeLearning, que permite agregar interactividad a videos de YouTube o de creación propia. Se puede agregar texto o imagen,

preguntas de respuesta única o múltiple, actividades de rellenar huecos, tareas desple-gables, emparejar elementos o lista desordenada.

Aquí, se presenta un análisis interpretativo de las intervenciones pedagógicas reali-zadas en los videos interactivos creados, a fin de identificar y caracterizar el tipo de interacción agregada en relación a la intencionalidad didáctica y la demanda cognitiva que genera. El análisis interpretativo de los datos buscó dar respuesta a las siguientes preguntas: ¿Qué tipo de intervenciones se incorporan en los videos? ¿Cuáles son las demandas cognitivas que requieren las intervenciones? y ¿Cuáles son las percepciones de los docentes respecto a las potencialidades de incorporar esta tecnología a sus prác-ticas?

La instancia formativa contó con la participación de veintitrés docentes (14 de Ma-temática, 8 de Física y 1 de Informática), distribuidos en catorce grupos de trabajo para el diseño de propuestas de enseñanza y Materiales Didácticos Digitales (MDD). Tras analizar las propuestas, se descartaron cinco por no incluir videos interactivos. Final-mente, de las propuestas consideradas se pudieron analizar 11 videos interactivos.

Para cada video incluido en el MDD se identifica la intención pedagógica, las inter-venciones realizadas y la fundamentación de las mismas. El análisis de las interven-ciones se organiza en torno a tres dimensiones: el tipo de pregunta, demanda cognitiva que genera y propósito pedagógico explicitado.

Los datos se sistematizaron en una tabla registrando en las filas los videos y en las columnas las características de las intervenciones. En la tabla 1 se presentan, a modo de ejemplo, los datos registrados para tres videos con sus correspondientes interven-ciones.

A partir de la tipología de procesos cognitivos propuesta por Manrique [7] se elabo-ran cinco categorías de análisis para tipificar la demanda cognitiva de las tareas pro-puestas en cada intervención incluida en el video: *de atención y percepción* (requieren focalizar la atención y filtrar estímulos relevantes); *de comprensión* (implican integrar información verbal y visual, identificar ideas centrales e inferir relaciones); *de inte-gración y retención* (buscan integrar la nueva información con conocimientos previos y sostener su retención en la memoria de trabajo); *de análisis y razonamiento* (proponen comparar, inferir o resolver problemas a partir del material presentado); y metacogniti-vas (invitan a reflexionar sobre la propia comprensión y regular las estrategias de apren-dizaje).

Al finalizar la propuesta de enseñanza, los docentes respondieron un cuestionario, diseñado y administrado mediante Google Forms, cuyo propósito fue conocer sus per-cepciones, experiencias y procesos de meta-reflexión en torno a la creación e inte-gración de videos educativos interactivos en sus propuestas de enseñanza. Se obtuvo información sobre: el perfil del docente (área disciplinar, nivel educativo y experiencia docente); su experiencia previa con videos interactivos; la potencialidad que perciben sobre el uso de videos interactivos; intenciones didácticas para incluirlos en las pro-puestas y las estrategias de diseño que considera más efectivas; percepción sobre conocimientos, destrezas y expectativas para diseñarlos y finalmente sobre la experi-encia de formación.

Tabla 1. presenta nro. de propuesta, nro. de video, contenido disciplinar y herramienta; momento de cada intervención expresado en minutos y segundos; tipo de fotograma (TX=texto; RU=respuesta única; RM=respuesta múltiple; RH=rellenar huecos; AD=actividad desplegable; EE=emparejar elementos; LD=lista desordenada; PA=pregunta abierta) y actividad propuesta.

Propuesta / video/contenido /Herramienta	t	Fotogramas								Actividad propuesta en la intervención
		TX	RU	RM	RH	AD	EE	LD	PA	
P1/ V1 / Concepto de función/eXeLearning	1:20		X							Responder a una pregunta dicotómica si la relación representa una función.
	4:01				X					Completar los conceptos claves en la definición de función.
	5:22	X								Reconocer los conceptos relacionados con función mediante el juego ruleta de palabras
P1/ V2/ Función lineal/eXeLearning	1:37				X					Completar una oración que se refiere metafóricamente a las características de la función lineal.
	2:36			X						Reconocer la definición correcta de pendiente.
	4:08			X						Seleccionar el método correcto para representar gráficamente la función.
P10 /V1 /Conceptos básico Estequiometría/eXelearning	1:41			X						Aplicar las nociones descritas en el video en el análisis de situaciones cotidianas
	1:51		X							Elegir la definición de mol que considere correcta de acuerdo a sus conocimientos previos (anticipación).
	3:42		X							Elegir la definición de masa molar que considere correcta según sus conocimientos previos (anticipación).

5:22	X	Calcular la masa molar de un compuesto con la información proporcionada en el video
7:59	X	Aplicar los conceptos descriptos en el video en situaciones sencillas. Ej.: ¿Cuántas partículas hay contenidas en 35 g de CH ₄ ?
8:15	X	Leer una síntesis de los conceptos presentados en el video.

4 Intervenciones Pedagógicas en videos interactivos

El análisis de los datos revela diseños instruccionales que aprovechan el potencial de los vídeos interactivos para transformarlos en herramientas de anclaje conceptual. De este modo, se fomenta el aprendizaje activo y se facilita la comprensión de los contenidos.

En lo que refiere al *tipo de intervención*, con eXelearning predominan las preguntas de Respuesta única o de selección múltiple y las actividades de rellenar huecos. En la fundamentación, que acompaña a una de las de las propuestas, mencionan que “*la herramienta (eXelearning) permite estructurar la secuencia de actividades, organizar los recursos y generar un material navegable que los estudiantes pueden explorar a su propio ritmo*”. El video, intervenido mediante la herramienta Edpuzzle, se emplea en una actividad de repaso de conceptos previos. En él se integraron preguntas abiertas que requieren, por ejemplo, identificar puntos notables de la función cuadrática o resolver una actividad en lápiz y papel y comparar con la resolución que muestra el video o responder preguntas en base a la solución obtenida (por ej. construir un gráfico). Asimismo, otra de las intervenciones incluye opciones de respuesta múltiple para identificar valores característicos en las gráficas, brindando la retroalimentación correspondiente.

Respecto a la *intencionalidad*, los docentes manifiestan que utilizan los videos principalmente para indagar conocimientos previos, recuperar y construir nuevo conocimiento y fomentar la autonomía en el aprendizaje. La principal estrategia identificada en las intervenciones es la activación y anticipación, cuyo propósito es indagar en los conocimientos previos al exigir que el espectador procese la información antes de recibir la respuesta. En este contexto, los vídeos interactivos resultan idóneos para presentar contenidos de forma sencilla, facilitando la recuperación de esquemas mentales y la generación de anticipaciones. Al abrir “estantes” en la memoria para alojar la nueva información, se reduce la carga cognitiva y se establece un diálogo efectivo entre los saberes previos y los nuevos contenidos que el video presentará tras la intervención. Por ejemplo, en una de las propuestas los docentes-cursantes manifiestan que los videos

interactivos les permite incluir *“preguntas para despertar el interés y favorecer la observación de la situación planteada”*; también mencionan que pueden *“indagar los conocimientos previos y establecer vínculos entre el contenido matemático y situaciones habituales de la vida cotidiana.”* En otra propuesta, expresan que *“con el video interactivo, se pretende retomar conceptos anteriores y también reforzar y resignificar los nuevos conceptos”*.

Además, se identifica una estrategia de combinar lo visual y auditivo del video con textos para favorecer la comprensión de los conceptos más importantes, como por ejemplo notas aclaratorias, opciones para elegir la definición correcta entre varias propuestas, opciones para completar una proposición, o síntesis de los conocimientos nuevos presentados en el video. Expresan que la intencionalidad de las intervenciones es, por ejemplo, *“retomar conceptos anteriores, reforzar y resignificar los nuevos conceptos”*; o como *“repaso y profundización de los contenidos estudiados”*.

En algunos casos, se observa que las proposiciones carecen de precisión lo que podría obstaculizar la interpretación del texto y comprometer la comprensión de los conceptos.

Finalmente se reconoce, en las intervenciones agregadas, estrategias de transferencia y aplicación como por ejemplo cálculos asistidos donde el video sirve de modelo; análisis de situaciones cotidianas; o aplicación de fórmulas. En relación a esta estrategia en una de las propuestas mencionan que su intencionalidad es que: *“Durante la visualización, se incorporan preguntas interactivas que estimulan la reflexión y la interpretación intuitiva de lo que sucede. Este tipo de recursos promueve la autonomía del aprendizaje, ya que las respuestas reciben retroalimentación inmediata, permitiendo a los estudiantes verificar sus interpretaciones y construir progresivamente una comprensión significativa del fenómeno físico”*.

4.1 Demandas Cognitivas

Basándonos en las fundamentaciones teóricas respecto al tipo de demanda cognitiva y al contenido didáctico-disciplinar de cada intervención, se caracterizan y ejemplifican las demandas cognitivas identificadas en las propuestas.

Demandas de Atención y Percepción: Este nivel de demanda se reconoce en todos los casos, el diseño interactivo busca reorientar la atención del estudiante mediante interrupciones estratégicas. El objetivo explícito es *“usar preguntas para despertar el interés y favorecer la observación de la situación planteada”*. Al plantear situaciones visuales (como el movimiento de un émbolo o gráficos de funciones) y detener el video, se demanda al estudiante una percepción selectiva para identificar elementos visuales clave antes de continuar.

Demandas de Comprensión: El diseño de la interacción exige que el estudiante traduzca el lenguaje audiovisual a conceptos abstractos. Por ejemplo: en matemática, se busca establecer vínculos entre el contenido matemático y *“situaciones habituales de la vida cotidiana”*. Las intervenciones no solo piden repetir datos, sino entender el significado de lo que se ve. Por ejemplo, en Física relacionar la presión con el volumen, en programación se explora explícitamente si se comprendió el uso de los distintos elementos de las estructuras, ya sea en relación con la sintaxis o con la semántica.

Demandas de *Integración y Retención*: Esta demanda predomina en las propuestas de matemática. Se utilizan actividades para recuperar conocimientos previos y reforzar y resignificar los nuevos conceptos. Específicamente, se incluyen actividades de rellenar huecos, la que es una tarea clásica de recuperación de información. En estos casos el video funciona como un nodo integrador que conecta lo que el estudiante ya sabía con lo nuevo. La demanda cognitiva aquí es recuperar información de la memoria a largo plazo para integrar con lo que presenta el video.

Demandas de *Análisis y Razonamiento*: Aunque muchas preguntas son cerradas, la intencionalidad descrita apunta a procesos de pensamiento superior, especialmente en la enseñanza de Física. En estos casos las preguntas estimulan la reflexión y la interpretación de lo que sucede. Se solicita el análisis de un fenómeno, por ejemplo *¿qué pasa con la presión al apretar el émbolo de una jeringa?* y luego deducir una regla o principio. Esto supera la simple memorización dado que exige inferencia lógica.

Demandas *Metacognitivas*: Esta categoría se evidencia en la fundamentación de los docentes, quienes ven al video como herramienta de autorregulación. Valoran que los recursos sean autoevaluativos y con retroalimentación inmediata, *que "permite a los estudiantes verificar sus interpretaciones y construir comprensión de manera progresiva"*. Al recibir feedback inmediato, por ejemplo cuándo se pregunta *¿cómo te quedó el gráfico: bien o mal?* el estudiante se ve forzado a monitorear su propio proceso de aprendizaje, activando la conciencia sobre el propio error y la corrección en tiempo real.

El análisis de los datos revela que los videos interactivos propuestos por los docentes configuran un entorno de aprendizaje activo caracterizado por una fuerte orientación hacia la atención, la retención y la metacognición. Han sido diseñados priorizando dos ejes cognitivos: la integración y retención (vincular conocimientos previos con nuevos) y la regulación metacognitiva (uso del error y el feedback inmediato para la autoevaluación). Se observa el anclaje atencional como base del diseño, una demanda alta en cuanto a atención y percepción. A la vez existe una demanda alta de Integración y retención, estrategia centrada en la construcción y retención del conocimiento (vincular conocimientos previos con nuevos). En un nivel intermedio, se sitúan las demandas de Comprensión y Análisis/Razonamiento. Si bien se busca la "comprensión conceptual" y "establecer vínculos con situaciones de la vida real", el razonamiento suele estar orientado por la estructura cerrada de las preguntas. Finalmente, se observa una demanda más alta de habilidades Metacognitivas. Los docentes fundamentan el uso de estas herramientas en su capacidad para fomentar la "autonomía en el aprendizaje" mediante recursos "autoevaluativos".

En síntesis, la narrativa que surge de los datos muestra un diseño centrado en la fijación de contenidos y la autorregulación, donde el estudiante es constantemente desafiado a prestar atención, recordar lo aprendido y evaluarse a sí mismo, aunque con un margen más acotado para el análisis exploratorio libre.

A modo de ejemplo, se muestran dos intervenciones en videos que acompañan una propuesta de enseñanza sobre la ley de los gases ideales. La **Figura 1** muestra una imagen de la intervención agregada al video que implica una demanda de *Integración y Retención*, dado que recupera la noción matemática de proporcionalidad y la resignifica en la relación entre presión y volumen del aire atrapado en la jeringa. Mientras,

la **Figura 2** muestra una imagen de una intervención que *demanda Análisis y Razonamiento*, dado que la pregunta estimula la reflexión y la interpretación de lo que sucede.



Figura 1: ejemplo de intervención de integración.



Figura 2: ejemplo de intervención de análisis.

Las producciones muestran que las intervenciones agregadas a los videos interactivos generan demandas cognitivas diversas. Si bien predominan aquellas vinculadas con la atención, la integración conceptual y la metacognición, también se identifican intervenciones que requieren análisis y razonamiento, especialmente en propuestas del área de Física. Esto sugiere un avance en la incorporación de tareas de mayor complejidad cognitiva, aun cuando persisten tendencias hacia actividades más estructuradas o de recuperación de información.

4.2 Percepciones, experiencias y metareflexión de los docentes

El análisis del cuestionario sobre la utilización y el diseño de videos educativos interactivos combina datos cuantitativos y cualitativos con el objeto de identificar fortalezas, barreras y oportunidades para la incorporación de este tipo de recursos en la práctica educativa. Participaron voluntariamente trece docentes del curso, provenientes de distintos niveles y áreas disciplinares, con experiencia profesional de más de diez años mayoritariamente.

Los resultados muestran que, aunque la mitad no había utilizado videos interactivos antes de la diplomatura, dan una alta valoración de su utilidad pedagógica. Los participantes destacan como principales beneficios la promoción del aprendizaje activo, la mejora de la atención, la retroalimentación inmediata y la integración de recursos multimedia.

Las principales limitaciones identificadas refieren a la falta de tiempo para el diseño, las necesidades de adquirir mayor capacitación técnica y las dificultades accesibilidad a la tecnología o de conectividad de los estudiantes. Los docentes señalan como principales necesidades formativas el fortalecimiento de estrategias didácticas para integrar videos interactivos.

Las interacciones más valoradas incluyen preguntas de opción múltiple, rellenar huecos y la retroalimentación automática.

En cuanto a la intención de uso futuro, la mayoría afirma que incorporaría videos interactivos con frecuencia, especialmente para actividades de evaluación formativa, revisión de contenidos, introducción de temas y trabajo autónomo. En el siguiente gráfico (ver **Figura 3**) se muestra las preferencias de uso de videos interactivos manifestada por los docentes en actividades de aprendizaje.

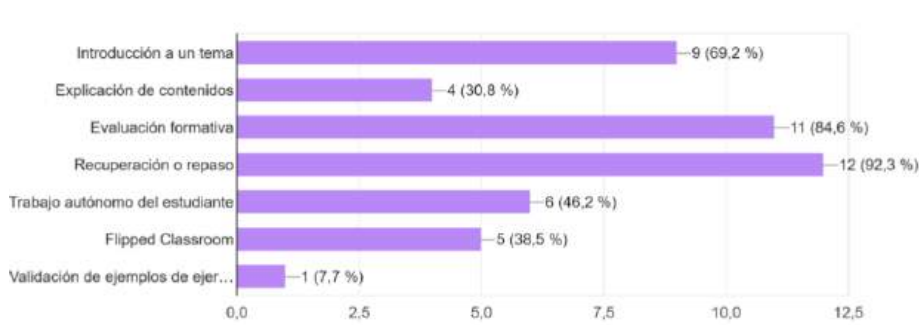


Figura 3. Intensión de uso de videos interactivos

Finalmente, la experiencia en el diseño de videos interactivos es percibida como enriquecedora, desafiante y útil para reflexionar sobre la propia práctica docente. Puede concluirse que los datos evidencian una alta predisposición a integrar videos interactivos en la enseñanza. En palabras de los docentes:

“La experiencia de diseñar el video educativo interactivo fue sumamente enriquecedora. Me permitió profundizar en cómo los materiales digitales pueden potenciar la comprensión y transformar la forma en que enseñamos y aprendemos”

“analizar en qué momentos del video incluir cada pregunta me permitió reflexionar sobre la secuencia didáctica y tomar decisiones más intencionadas. Fue un proceso productivo y me ayudó a profundizar en el diseño de recursos educativos interactivos.”

Los docentes valoran el potencial pedagógico de los videos interactivos para mejorar la atención, promover el aprendizaje significativo y proporcionar retroalimentación

inmediata. Las percepciones recogidas se encuentran alineadas con los resultados obtenidos previamente en relación con la intencionalidad pedagógica.

Los videos interactivos se integraron de forma coherente con los contenidos y la intencionalidad didáctica. El análisis demuestra que los docentes lograron reconceptualizar el video como recurso didáctico: pasaron de un uso pasivo a un diseño que favorece el aprendizaje activo y autónomo.

5 Comentarios finales

El estudio se centra en el diseño de videos interactivos por parte de un grupo reducido de veintitres docentes de áreas de las ciencias exactas (Matemática, Física e Informática), cuyas producciones fueron el resultado de un proceso de acompañamiento académico orientado.

El diseño de estos videos consistió en la incorporación de intervenciones mediante herramientas como eXeLearning y Edpuzzle, con el propósito de transformar videos disponibles en la web en recursos interactivos. Se observó que los docentes recurren principalmente a estrategias de memorización y recuperación de la información, evidenciado en el uso predominante de intervenciones basadas en preguntas de opción múltiple y ejercicios de completar. Estas herramientas se emplean, en particular, para la activación de conocimientos previos y la integración de nueva información, funcionando como anclajes conceptuales que contribuyen a reducir la sobrecarga de la memoria de trabajo.

Asimismo, se evidencia que la planificación de intervenciones para el diseño de tareas interactivas, en función de las demandas cognitivas, favorece procesos reflexivos en los docentes sobre su propia práctica. Este proceso implica la articulación entre el contenido disciplinar y el material multimedia, y contribuye a ampliar el repertorio de estrategias pedagógicas, promoviendo la diversificación y el enriquecimiento de los aprendizajes mediados por tecnologías. En este sentido, el diseño pone de manifiesto la necesaria integración de saberes tecnológicos, pedagógicos y disciplinares.

Por otra parte, los participantes valoraron positivamente el uso didáctico de los videos interactivos, reconociéndose como recursos enriquecedores y versátiles, especialmente adecuados para la evaluación formativa y el repaso de contenidos. Los consideran una oportunidad para implementar estrategias de enseñanza orientadas al aprendizaje activo. No obstante, para lograr una implementación sostenible, resulta necesario abordar ciertas barreras, tales como la limitación de tiempo para el diseño y la necesidad de una mayor capacitación técnica.

Referencias

1. Santos, G. Miranda, A. y Abásolo Guerrero, M. J., (2021). Intervenciones pedagógicas para crear videos interactivos, *Proceedings of X Ibero-American Conference on Applications and Usability of TVDI jAUTI2021*. <https://doi.org/10.35537/10915/143302>
2. Vallejos, A. y González, A., (2018). Diseño de videos más efectivos, enriquecidos con interactividad, utilizando la herramienta H5. En Actas del Vº Foro Internacional de Educación en Entornos Virtuales y XVIIIº encuentro AIESAD. Universidad Nacional de Quilmes, Quilmes. <http://sedici.unlp.edu.ar/handle/10915/69685>

3. Mayer, R.E. (2024) The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educ Psychol Rev* 36, 8. <https://doi.org/10.1007/s10648-023-09842-1>
4. Asunción, S. (2019). Metodologías Activas: Herramientas para el empoderamiento docente. *Revista Docentes 2.0*, 7(1), 65–80. <https://doi.org/10.37843/rted.v7i1.27>
5. Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.004>
6. Benedicto, C., Jaime, A. y Gutiérrez, Á. (2015). Análisis de la demanda cognitiva de problemas de patrones geométricos. En C. Fernández, M. Molina y N. Planas (Eds.), *Investigación en Educación Matemática XIX* (pp. 153-162). Alicante: SEIEM.
7. Manrique, M. S. (2020). Tipología de procesos cognitivos. Una herramienta para el análisis de situaciones de enseñanza. *Educación*, 29(57), 163–185. <https://doi.org/10.18800/educacion.202002.008>
8. Harris, J., Mishra, P. & Koehler, M. (2009). Teachers' Technological Pedagogical Content Knowledge and Learning Activity Types: Curriculum-Based Technology Integration Reframed. *Journal of Research on Technology in Education*, 41(4), 393-416. Retrieved December 13, 2025 from <https://www.learntechlib.org/p/106092/>



Each memory unfolds a new story

This workshop aims to bring together researchers, practitioners, educators, and creative professionals who explore storytelling practices in digital, interactive, and immersive contexts. The sessions are promoted by Polariscope, a research and development (R&D) project focused on collaborative storytelling through media content. The Polariscope platform seeks to bring citizens and institutions together to share and interact with memories related to cultural heritage, significant events and experiences, combining multiple perspectives.

<https://polariscope.web.ua.pt/>



polariscope WORKSHOP

Digital Storytelling and AI in Co-Creation processes

Digital Storytelling for Student Engagement



Para Além dos Pontos e Distintivos: Desenhar a Motivação Intrínseca nas Ferramentas Digitais de SEL

Raquel Magalhães¹

¹Departamento de Comunicação e Artes, Digimedia Research Centre,
Universidade de Aveiro, Portugal
raquel.magalhaes@ua.pt

Sumário. Este estudo examina o papel da motivação no design de ferramentas digitais voltadas à Aprendizagem Social e Emocional (SEL) para estudantes do ensino secundário. O artigo analisa especificamente a dimensão motivacional do modelo SEED.IC (Social Emotional Educational Digital Infocommunication Model), desenvolvido por meio de uma abordagem de Design Based Research envolvendo 149 participantes em Portugal e Espanha. A investigação focaliza três características motivacionais centrais: gamificação, simulações baseadas em role playing e incentivos contextuais (nudges). Os resultados revelam percepções divergentes entre stakeholders, evidenciando que estratégias baseadas exclusivamente em recompensas extrínsecas podem comprometer a motivação intrínseca dos adolescentes. Em contraste, abordagens que privilegiam autonomia, reconhecimento entre pares e experiências narrativas demonstram maior alinhamento com as necessidades de desenvolvimento socioemocional. Conclui-se que mecanismos motivacionais eficazes em contextos digitais de SEL devem estar estruturalmente ancorados em fundamentos éticos e sociais, priorizando o crescimento intrínseco e a autogestão dos estudantes.

Palavras-chave: aprendizagem social e emocional; motivação intrínseca; gamificação; tecnologia educacional; SEED.IC.

1 Introdução

O crescente reconhecimento da Aprendizagem Social e Emocional (SEL) como componente crítico da educação contemporânea (OCDE, 2018, 2025) exige o desenvolvimento de intervenções digitais estruturadas e baseadas em evidências que apoiem adolescentes em desenvolvimento (Huttunen et al., 2024; e Yeager, 2017). Embora os programas tradicionais de SEL tenham dependido predominantemente de métodos de entrega presenciais (CASEL, 2025), a rápida digitalização dos ambientes educativos (Vuorikari et al., 2022) exige estruturas que possam orientar o design eficaz de soluções de SEL melhoradas tecnologicamente (Erstad et al., 2024). O modelo de Informação e Comunicação Digital Social e Emocional Educativa (Social Emotional Educational Digital Infocommunication Model, com sigla SEED. IC em inglês) surgiu para colmatar esta lacuna, fornecendo um quadro abrangente para a criação de artefactos digitais especificamente

adaptados aos contextos do ensino secundário (Magalhães, 2025).

O modelo SEED.IC compreende 18 características distintas de comunicação de informação, organizadas em seis dimensões: social, emocional, ética, motivacional, educacional e técnica. Dentre estas, a dimensão motivacional representa um conjunto crítico de funcionalidades concebidas para captar e manter o envolvimento dos utilizadores, tornando-se um determinante chave do impacto a longo prazo e do sucesso educativo do artefacto digital. Esta dimensão inclui características especificamente focadas em manter a interação, especialmente do público de adolescentes, com as ferramentas digitais de SEL através de estratégias de envolvimento significativas.

No entanto, as partes interessadas envolvidas no desenvolvimento do modelo enfatizaram um princípio crucial de design: as características motivacionais não podem funcionar isoladamente, mas devem ser estruturalmente apoiadas por características éticas e sociais fundamentais. Este artigo examina especificamente a criação conceptual, validação empírica e desempenho contestado de três características motivacionais (Gamificação, Simulações-Role Playing, e Incentivos Contextuais, ou Nudges) argumentando que o envolvimento bem-sucedido depende fundamentalmente de priorizar o crescimento intrínseco e a autonomia do aprendiz em detrimento dos mecanismos extrínsecos de recompensa.

2 Revisão da Literatura

O design de funcionalidades motivacionais na tecnologia educativa tem sido dominado por abordagens de gamificação que enfatizam recompensas extrínsecas como pontos, distintivos e tabelas de classificação (Montiel-Ruiz, Sánchez-Vera & Solano-Fernández, 2023). Embora estes mecanismos possam gerar envolvimento a curto prazo, a investigação sugere cada vez mais que podem minar a motivação intrínseca, especialmente entre adolescentes que estão a desenvolver autonomia e capacidades de pensamento crítico. A Teoria da Autodeterminação (Deci & Ryan, 2000) aponta que a motivação sustentável surge da satisfação de três necessidades psicológicas fundamentais: autonomia, competência e relação. Intervenções digitais que priorizam recompensas externas correm o risco de comprometer estas necessidades, podendo levar ao “desengajamento” dos utilizadores quando os efeitos de novidade diminuem.

No contexto da SEL, esta tensão torna-se particularmente pronunciada. As competências sociais e emocionais, incluindo autoconsciência, autogestão, consciência social, competências relacionais e tomada de decisões responsáveis, são processos inerentemente de desenvolvimento que requerem autorreflexão genuína e mudança comportamental autêntica (CASEL, 2025). Desenhos motivacionais que enfatizam o envolvimento superficial ou táticas manipuladoras contradizem os objetivos fundamentais da educação SEL, que procuram cultivar indivíduos emocionalmente inteligentes e autónomos, capazes de navegar em ambientes sociais complexos (Durlak et al., 2015).

A dimensão motivacional do modelo de SEED.IC foi, por isso, conceptualizada

em torno de três características centrais concebidas para equilibrar o envolvimento com a adequação ao desenvolvimento:

1. Gamificação: A aplicação estratégica dos princípios do design de jogos em contextos educativos não recreativos, indo além dos simples sistemas de pontos em direção à visualização significativa do progresso e reconhecimento de conquistas.
2. Simulações - Role Playing: Capacidades interativas que permitem aos utilizadores praticar competências sociais e emocionais em cenários seguros e narrativos que refletem desafios sociais do mundo real.
3. Incentivos Contextuais e Gentis ou *Gentle Nudges*: Prompts relevantes e oportunos, concebidos para orientar o comportamento do utilizador para funcionalidades e práticas benéficas sem coerção, respeitando a agência do utilizador enquanto facilitam a formação de hábitos.

Esta filosofia de design prescritiva reflete um mandato fundamental: as estratégias de envolvimento devem ser intencionalmente elaboradas para promover o crescimento intrínseco, em vez da manipulação ou superficialidade. O quadro teórico posiciona a motivação não como um fim em si, mas como um meio de apoiar a trajetória de desenvolvimento do adolescente rumo à autogestão e à competência emocional. E isto está alinhado com o apelo prático do SEL para ser significativo para cada público (Elias, 2014).

3 Método

O modelo de SEED.IC, incluindo a sua dimensão motivacional, foi realizado através da Investigação Baseada em Design (DBR), baseada em Wang & Hannafin (2005), envolvendo 149 participantes em Portugal e Espanha, e buscou a ampliação de experiências para uma conclusão participativa ao invés de somente um consenso de opiniões. Esta abordagem metodológica permitiu o refinamento iterativo do modelo com base na contribuição das partes interessadas e testes empíricos, garantindo validade ecológica e aplicabilidade prática. O processo de investigação incorporou múltiplos métodos complementares:

Validação Quantitativa: Todos os participantes avaliaram as 18 características do SEED.IC usando uma escala de Likert de cinco pontos (1 = não importante; 5 = extremamente importante). Esta avaliação gerou classificações de importância comparáveis em todas as dimensões do modelo, permitindo a identificação de áreas de consenso, oportunidades de reflexão e características contestadas.

Atividades do Mapa de Calor da SEL: Foram atividades estruturadas onde distribuíram votos entre os cinco domínios de competência da SEL (autoconsciência, autogestão, consciência social, competências relacionais, tomada de decisões responsáveis). Este método revelou as prioridades e necessidades percebidas pelos alunos, professores e experts informando o alinhamento das

características motivacionais com preocupações autênticas de desenvolvimento.

Ordenação de Cartões com Especialistas, Alunos e Professores: Exercícios de triagem de cartões para categorizar e priorizar as 18 características dentro de dimensões as quais os participantes criavam nomes ou categorias. Estas sessões geraram dados qualitativos e entre as discussões podemos notar as interdependências entre dimensões e os potenciais riscos associados a funcionalidades motivacionais mal implementadas.

Testes Iterativos de Protótipos: A aplicação SUNRISE, desenvolvida como uma instância prática do SEED. IC, passou por ciclos de testes com alunos, experts e professores do ensino secundário. Foi recolhido feedback direto sobre funcionalidades motivacionais, particularmente elementos de gamificação, através de entrevistas e observações de usabilidade, permitindo ajustes de design em tempo real.

Esta abordagem multimétodo proporcionou tanto amplitude (avaliações quantitativas de importância) como profundidade (explicações qualitativas de preferências e preocupações), gerando uma compreensão abrangente de como as partes interessadas percebem e valorizam diferentes estratégias motivacionais.

4 Principais reflexões

A avaliação empírica da dimensão Motivacional revelou divergências significativas nas perceções dos stakeholders, destacando a complexidade e contestação em torno das características orientadas para o envolvimento no design digital de SEL, conforme pode ser visto na Figura 1.

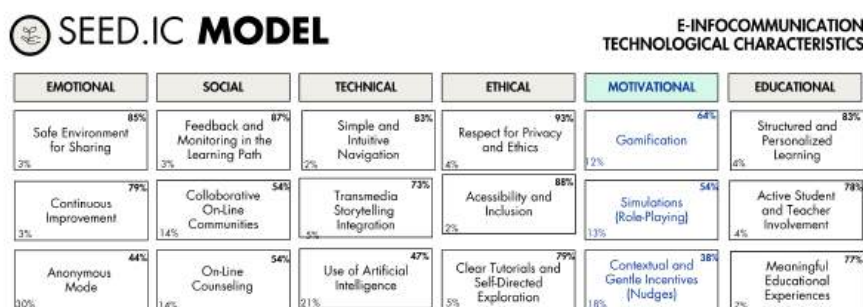


Figura 1. Modelo SEED.IC com destaque das características motivacionais

4.1 Métricas Quantitativas de Desempenho

A "gamificação" demonstrou o maior apoio entre as três características motivacionais, com 64% dos participantes a considerarem altamente importante (pontuações 4 ou 5 na escala de Likert). No entanto, este apoio aparentemente forte escondeu uma polarização considerável: 12% dos participantes atribuíram classificações de baixa importância (pontuações 1 ou 2), assinalando a gamificação como uma característica com "perceções mistas" e substancialmente menos consenso em comparação com elementos fundamentais do modelo.

Esta divergência torna-se particularmente notável quando contrastada com o acordo quase universal sobre fundamentos éticos. "Respeito pela Privacidade e Ética", posicionado na estrutura do modelo, recebeu avaliações de 93% de alta importância, tornando-se a característica mais consensual em todo o enquadramento. Esta diferença entre a gamificação e as fundações éticas demonstra empiricamente que, embora as características de envolvimento sejam valorizadas, são vistas como significativamente mais contestadas e condicionais do que os elementos de construção de confiança.

"Simulations - Role Playing" recebeu 54% de classificações de alta importância, mas foi oficialmente categorizado como "divisivo" com base nos seus 13% de baixa importância. Os dados qualitativos sugeriram que esta divisão não derivava da oposição à aprendizagem baseada em simulação em si, mas de preocupações quanto à qualidade da implementação e adequação dos cenários.

"Incentivos Contextuais e Gêntis - Empurrões" apresentou o desempenho mais fraco entre todas as características motivacionais, recebendo apenas 38% de avaliações de alta importância, a pontuação mais baixa entre o trio motivacional. Além disso, esta funcionalidade recebeu 18% de avaliações de baixa importância, sugerindo "baixa priorização e ceticismo" quanto à sua utilidade em contextos de SEL. Os participantes questionaram se os mecanismos de empurrão poderiam minar a autonomia dos adolescentes ou ser vistos como paternalistas.

4.2 Insights Qualitativos: A Tensão de Design

O feedback qualitativo forneceu um contexto crucial para compreender os resultados quantitativos mistos, revelando uma tensão fundamental de design entre imperativos de envolvimento e respeito pela maturidade adolescente.

Os alunos rejeitaram explicitamente sistemas de recompensa que consideravam infantis ou manipuladores. Durante os testes de protótipos da aplicação SUNRISE, as implementações iniciais de sistemas de gamificação baseadas em pontos foram criticadas pelos alunos por parecerem "truques" concebidos para os fazer usar a aplicação, em vez de ferramentas genuínas para autoaperfeiçoamento. Este feedback alinhou-se com os avisos dos educadores durante as sessões de triagem de cartões de que as soluções digitais SEL "só podem dar frutos se estiverem enraizadas em fundamentos emocionais e sociais e sustentadas por força ética e técnica."

Por outro lado, os alunos responderam positivamente a abordagens motivacionais baseadas no reconhecimento dos pares e no envolvimento narrativo. Cenários de role-playing baseados em simulação que lhes permitiram explorar

dilemas sociais e praticar estratégias de regulação emocional foram descritos como valiosos e respeitadores da sua inteligência. Estas descobertas reforçaram a premissa teórica de que os adolescentes priorizam a autonomia e a autenticidade em detrimento do entretenimento gamificado.

As atividades do SEL Heat Map forneceram uma perspetiva adicional sobre as prioridades motivacionais dos alunos. As competências de autogestão receberam 26% de todos os votos dos alunos, a maior alocação entre os cinco domínios SEL, indicando um forte "imperativo de autogestão" entre os adolescentes. Os alunos queriam ferramentas que os ajudassem a regular as emoções, definir e alcançar objetivos, e desenvolver hábitos construtivos. Esta descoberta sugeriu que as características motivacionais alinhadas com a autogestão (como o acompanhamento do progresso e os estímulos de reflexão) ressoariam de forma mais autêntica do que os sistemas genéricos de recompensa.

Com base neste feedback, os designers modificaram iterativamente a arquitetura motivacional da aplicação SUNRISE. Os sistemas iniciais de pontos foram substituídos por mecanismos de reconhecimento entre pares e estruturas de desempenho ligadas ao desenvolvimento genuíno de competências, em vez de métricas arbitrárias. Esta mudança demonstrou a aplicação prática do princípio prescritivo do modelo: a motivação deve servir ao crescimento intrínseco e não ao envolvimento superficial.

5 Conclusões

A dimensão motivacional do modelo SEED.IC demonstra uma tensão empírica crítica no design digital de SEL: embora os mecanismos de envolvimento sejam desejados pelos programadores e valorizados por algumas partes interessadas, abordagens simplistas de gamificação correm o risco de alienar os próprios utilizadores que pretendem servir, particularmente adolescentes que priorizam autonomia, maturidade e autenticidade.

Os dados indicam fortemente que o papel da motivação deve ser fundamentalmente redefinido nos contextos digitais de SEL. A "Coroa" Motivacional não pode ser isolada, mas está estruturalmente dependente das bases éticas e sociais representadas pelo "Tronco" e pelas "Raízes" do modelo. Sem confiança estabelecida, manifestada através de proteções rigorosas de privacidade, práticas éticas de dados e funcionalidade transparente, as características motivacionais são percebidas como manipuladoras em vez de apoiadoras. O consenso de 93% sobre fundamentos éticos contra 64% de apoio à gamificação valida empiricamente esta dependência hierárquica.

Para designers e investigadores em tecnologia educativa, estas descobertas têm implicações importantes. Um design motivacional eficaz deve alinhar explicitamente as estruturas de incentivos com os objetivos centrais de desenvolvimento, afastando-se decisivamente dos sistemas extrínsecos de recompensa (pontos, distintivos, tabelas de classificação) para mecanismos que

promovam o reconhecimento dos pares, proporcionem prática narrativa através de simulações realistas e, em última análise, cultivem competências intrínsecas de autogestão. Esta mudança assegura que a tecnologia cumpre a sua função adequada: apoiar o desenvolvimento de cidadãos digitais resilientes, emocionalmente inteligentes, capazes de uma autêntica autorreflexão e mudança comportamental, em vez de apenas captar a atenção através de entretenimento superficial ou padrões de design exploradores.

A abordagem prescritiva do modelo de IC, posicionando a motivação como contingente, contestada e fundamentalmente dependente de compromissos éticos mais profundos, oferece um roteiro para criar intervenções digitais SEL que respeitem o desenvolvimento adolescente mantendo um envolvimento significativo.

References

1. CASEL - Colaboração para Aprendizagem Académica, Social e Emocional. (2025). Fundamentos do SEL. <https://casel.org/overview-sel/>
2. Durlak, J. A., Domitrovich, C. E., Weissberg, R. P., & Gullotta, T. P. (2015). *Manual de Aprendizagem Social e Emocional: Investigação e Prática*. Guilford Press.
3. Elias, M. J. (2014). "O futuro da aprendizagem socioemocional: Um apelo à ação." *O Jornal de Psicologia Educacional*.
4. Erstad, O., Černochová, M., Knezek, G., Furuta, T., Takami, K., & Liang, C. (2024). Modos sociais e emocionais de aprendizagem em ecossistemas digitais: agendas emergentes de investigação. *Tecnologia, Conhecimento e Aprendizagem*, 29, 1751–1766. <https://doi.org/10.1007/s10758-024-09775-w>
5. Huttunen, I., Upadyaya, K., & Salmela-Aro, K. (2024). Associações longitudinais entre as competências socioemocionais dos adolescentes, o envolvimento escolar e o esgotamento escolar. *Aprendizagem e Diferenças Individuais*, 115, 102537. <https://doi.org/10.1016/j.lindif.2024.102537>
6. Magalhães, R. (2025) Um Modelo de Infocomunicação para a Educação Social e Emocional de Estudantes do Ensino Secundário: Um estudo multicaso em Aveiro e Barcelona. Tese de Doutoramento a ser apresentada em dezembro de 2025. Universidade de Aveiro.
7. Montiel-Ruiz, F. J., Sánchez-Vera, M. D. M., & Solano-Fernández, I. M. (2023). Redes sociais e gamificação na educação física: Um estudo de caso. *Tecnologia Educativa Contemporânea*, 15(1). <https://doi.org/10.30935/cedtech/12660>
8. OCDE. (2018). O Futuro da Educação e das Competências: Educação 2030. Documentos de Trabalho sobre Educação da OCDE, 23. https://www.oecd.org/content/dam/oecd/en/publications/reports/2018/06/the-future-of-education-and-skills_5424dd26/54ac7020-en.pdf
9. OCDE. (2025). OECD Teaching Compass: Uma ferramenta para apoiar a aprendizagem e desenvolvimento profissional dos professores. OECD Publishing. https://www.oecd.org/en/publications/oecd-teaching-compass_8297a24a-en.html
10. Ryan, R. M. & Deci, E. L. (2000). *Teoria da autodeterminação e a facilitação da motivação intrínseca, desenvolvimento social e bem-estar*. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
11. Vuorikari, R., Kluzer, S., & Punie, Y. (2022). DigComp 2.2: O Quadro de Competência Digital para Cidadãos - Com novos exemplos de conhecimento,

competências e atitudes. Gabinete de Publicações da União Europeia.
<https://doi.org/10.2760/115376>

11. Wang, F., & Hannafin, M. J. (2005). Investigação baseada em design e ambientes de aprendizagem melhorados por tecnologia. *Investigação e Desenvolvimento em Tecnologia Educacional*, 53(4), 5–23. <https://doi.org/10.1007/BF02504682>
12. Yeager, D. S. (2017). Programas de aprendizagem social e emocional para adolescentes. *O Futuro das Crianças*, 27(1), 73–94.
<https://files.eric.ed.gov/fulltext/EJ1145078.pdf>

Game Engagement and Environmental Literacy: Development and Validation of a Motivational Model in Digital Game Contexts

Mariana Pinho¹ [0000-0002-0173-444X], Enrickson Varsori^{2*} [0000-0002-8875-9278] and Abílio Almeida³ [0000-0003-0936-3898]

¹ ECOMARE, Centre for Environmental and Marine Studies, University of Aveiro, Aveiro, Portugal

² University of Aveiro, Communication and Society Research Centre, Institute of Social Sciences, University of Minho, Braga, Portugal, varsori@ua.pt

³ Communication and Society Research Centre, Institute of Social Sciences, University of Minho, Braga, Portugal

Abstract. The abstract should summarize the contents of the paper in short terms, i.e. 15–250 words. The relationship between engagement in digital games and pro-environmental attitudes has attracted growing attention in the field of interactive applications for literacy and sustainability. The research proposes a framework that brings together three core dimensions: Game Engagement, Motivation, and Environmental Attitudes (toward environmental preservation) to examine how game-based learning experiences can foster sustainable behaviors. Building on an initial instrument developed for young students and tested in a first phase with young people ($n = 124$), this research extends the theoretical model and prepares its psychometric validation using a new sample of students ($n = 114$). The questionnaire uses age-appropriate 5-point Likert scales with visual support (pictorial scale) to facilitate younger students' understanding of response intensity across the three constructs.

This paper outlines the conceptual and methodological foundations of the validation process, including the criteria for confirmatory factor analysis and the indicators planned to assess internal consistency, convergent validity, and measurement invariance across groups. The discussion emphasizes the potential of educational games as interactive learning environments capable of promoting motivation and engagement among younger generations, while integrating perspectives on usability, motivation, and social impact.

Keywords: Game Engagement; Motivation; Environmental Attitudes; Literacy Digital Games.

1 Introduction

Digital games have become increasingly common in classroom contexts over the last decade. Their appeal comes not only from their ability to capture students' attention at the start of an activity, but also from how they help maintain involvement as learners

progress through various tasks. Features such as interaction, immediate feedback, and simple narrative elements often make the experience more engaging for younger students [1, 2]. Research has shown that these characteristics create opportunities for exploration and decision-making, prompting emotional involvement, which in turn contributes to more meaningful learning [3, 4]. When applied to topics related to environmental and ocean literacy, these advantages become even more evident, as many abstract concepts gain clarity when illustrated through concrete and accessible examples for students [5–7]. If researchers aim to understand how students respond to such digital experiences, they require instruments capable of measuring key constructs with clarity and in language that is appropriate for younger learners. Several well-established tools exist, such as the Intrinsic Motivation Inventory [8–10], the Game Engagement Questionnaire [11], and scales assessing environmental attitudes [12, 13]. However, these instruments were originally designed for older populations or for contexts that differ considerably from short, narrative-oriented educational games. To address these limitations, an instrument was developed specifically for students in basic education, organized around three central dimensions: Motivation, Game Engagement, and Environmental Attitudes. While grounded in established theoretical frameworks, the items were adapted to ensure accessibility for younger learners and to reflect learning scenarios involving digital games focused on environmental themes. Although this instrument has been used in previous empirical studies, its psychometric properties have not been systematically examined. This article follows the common structure to validate studies: an Exploratory Factor Analysis (EFA) to examine how the items group together, followed by a Confirmatory Factor Analysis (CFA) to assess whether the proposed three-factor model provides a satisfactory fit. Confirming these psychometric properties ensures that the instrument can be reliably applied in studies investigating motivation, engagement, and environmental attitudes within digital game-based learning.

2 Methods

2.1 Study 1: Exploratory Factor Analysis (EFA)

Participants and Procedure. Data were collected from 124 students (61 male, 56 female, and seven who preferred not to disclose their gender). Participants ranged in age from 10 to 14 years ($M = 11.38$, $SD = 1.64$). Participants were recruited through advertisements in public schools in Portugal. Three 50-minute sessions were scheduled in the selected schools. The digital game used in the sessions was a short, narrative-driven educational activity focused on ocean and environmental themes. In each session, students engaged with the game in a classroom setting, and the gameplay experience served as the basis for the subsequent self-report assessment. Completing the questionnaire took, on average, 5 minutes. At the end of the session, participants were thanked and provided with a debriefing. Participation in the study was voluntary and anonymous. The study was conducted in accordance with the principles outlined in the Declaration of Helsinki. Ethical approval was not required for this study, as per the Resolution of the Assembly of the Portuguese Republic No. 29/2017, Articles 15 and 16. Informed consent was obtained from all participants and their legal guardians prior to

participation. Written consent for publication of the study's results was also obtained from both the participants and their legal guardians.

Measures. In all dimensions, responses were collected using 5-point Likert scales accompanied by visual support (pictorial scale) to help younger students better understand the relative weight of the response options.

Motivation. The motivation factor was based on the principles outlined in the Intrinsic Motivation Inventory (IMI) developed by [8] within the framework of Self-Determination Theory (SDT). Both intrinsic motivations (e.g., “*I enjoyed playing this game*”) and extrinsic motivations (e.g., “*I wanted to play more games about the ocean and nature*”) were measured. The scale was adapted in accordance with the theoretical guidelines established by [9, 10]. A subscale consisting of five statements was created to assess motivation, using a 5-point Likert scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*), regarding the digital game.

Game Engagement. The game engagement factor was developed by adapting the Game Engagement Questionnaire (GEQ) by [11], which assesses the extent of player involvement during gameplay (e.g., “*I felt challenged to learn more about the ocean and nature while playing*”). This dimension encompasses statements related to aspects such as focus, time perception, and immersion in the game. A subscale of five statements was created for this purpose, assessed using a 5-point Likert scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*), specifically related to the digital game experience.

Environmental Attitudes. The environmental attitudes factor was based on prior research on young people's pro-environmental attitudes, particularly the works of [12] and [13]. This measure was designed to capture the self-reported environmental attitudes of young people, particularly in relation to environmental preservation (e.g., “*Would you like to learn more about how to care for the ocean and nature?*”). The dimension was assessed using a 5-point Likert scale, with ratings ranging from 1 (*Not Important*) to 5 (*Very Important*).

Socio-Demographic Variables. Participants were asked to report their age, gender, and the school year in which they were currently enrolled.

Results.

Exploratory Factor Analysis. The exploratory factor analysis showed good overall model fit [14, 15]. The chi-square test was significant, $\chi^2(78) = 575, p < .001$. The Root Mean Square Error of Approximation (RMSEA) value of 0.060 indicated an acceptable level of approximate fit, and the TLI value of 0.93 suggested good model fit relative to the baseline model. The Bayesian Information Criterion (BIC = -140) further supported the adequacy of the model. All items demonstrated acceptable uniqueness values (all < 1), and the SS loadings were above 1 and statistically significant, indicating strong factor contributions. Sampling adequacy was confirmed by the KMO measure, with all MSA values exceeding the recommended threshold of 0.50.

Table 1. Factor Loadings and Cronbach's Alpha Reliabilities for Study Variables. MOTIV – Motivation; GE – Game Engagement; EA – Environmental Attitudes.

Variables	Items	Factor			Uniqueness	SS Loadings	% of Variance	Cumulative %	KMO MSA
		1	2	3					
MOTIV	MOTIV1	0.78			0.44	3.20	25	25	0.81
	MOTIV2	0.79			0.41				0.85
	MOTIV3	0.54			0.58				0.90
	MOTIV4	0.62			0.55				0.91
	MOTIV5	0.50			0.61				0.90
GE	GE1	0.44			0.69	1.65	13	38	0.88
	GE2			0.41	0.80				0.69
	GE3			0.92	0.14				0.83
	GE4	0.40		0.33	0.54				0.88
	GE5	0.71			0.41				0.87
EA	EA1		0.99		0.01	1.62	12	50	0.58
	EA2		0.32	0.38	0.71				0.71
	EA3		0.65		0.56				0.60

2.2 Study 2: Confirmatory Factor Analysis

Methods.

Participants. Data were collected from 238 (116 male, 108 female, and fourteen who preferred not to disclose their gender). Participants ranged in age from 10 to 14 years ($M = 12.16$, $SD = 1.45$).

Procedure and Measures. The same procedure and measures used in study 1 were applied. Due to Sample 2 being too small on its own ($n = 114$) to ensure stable and reliable parameter estimates, the confirmatory factor analysis (CFA) was conducted using the combined dataset from both samples. This approach is methodologically acceptable and commonly recommended when the exploratory factor analysis (EFA) has been performed on an independent sample. By merging the samples for the CFA, we increase statistical power, improve the precision of parameter estimation, and reduce the likelihood of improper solutions, while still preserving the independence between the EFA and CFA phases.

Results. Confirmatory Factor Analysis. A Confirmatory Factor Analysis was conducted, and the following results were obtained: $\chi^2(62) = 132$, $p < .001$; Goodness of Fit Index (GFI) = 0.99; Bentler-Bonett Normed Fit Index (NFI) = 0.99; Bollen's Relative Fix Index (RFI) = 0.98; Adjusted Goodness of Fit Index (AGFI) = 0.98; Tucker-Lewis Index (TLI) = 0.91; Comparative Fit Index (CFI) = 0.93; Root Mean Square Error of Approximation (RMSEA) = 0.07. Values for CFI, AGFI, GFI and TLI were greater than 0.9, indicating a good model fit (Cheung & Rensvold, 2002; Hu & Bentler, 1999). RMSEA was less than 0.08, which is considered an acceptable approximation error. Overall, the results showed a moderate to good model fit with the given data.

Additionally, convergent validity was assessed by the Average Variance Extracted (AVE). Most AVE values were higher than 0.5. One of the AVE values was not; however, composite reliability was higher than 0.6, indicating that the convergent validity of the constructs was acceptable [16]. Finally, discriminant validity was assessed by the Heterotrait-monotrait ratio of correlations (HTMT) as introduced by [17], addressing the limitations of the [16]. Most HTMT values were below 0.85, indicating good discriminant validity. The HTMT value between Motivation and Game Engagement (0.93) exceeded the commonly recommended threshold for discriminant validity. This result indicates a strong association between the two constructs. However, according to [17], higher HTMT values may occur when constructs are conceptually related or when participants perceive them as closely connected. This is particularly relevant in the context of game-based learning with younger students, where motivational and engagement responses often emerge simultaneously during play.

Previous research in the field of digital educational games supports this interpretation. Authors like [2] argue that motivation, cognitive engagement, and emotional involvement frequently co-occur in game-based environments, especially for children and adolescents, who tend to experience gameplay as a unified process rather than as distinct psychological dimensions. This overlap helps explain the strong correlation found in our data. Given that the literature supports the theoretical structure of the instrument and that all other psychometric indicators (model fit, factor loadings, reliability, and convergent validity) were satisfactory, the three-factor model was retained.

Table 2. Factor Loadings and Cronbach's Alpha Reliabilities for Study Variables. MOTIV – Motivation; GE – Game Engagement; EA – Environmental Attitudes

Variables	Items	Factor loadings	β	Cronbach alpha	Composite Reliability	Average Variance Extracted
MOTIV	MOTIV1	0.47	0.82	0.81	0.82	0.60
	MOTIV2	0.60	0.87			
	MOTIV3	0.49	0.75			
	MOTIV4	0.51	0.64			
	MOTIV5	0.49	0.76			
GE	GE1	0.52	0.62	0.77	0.77	0.48
	GE2	0.50	0.58			
	GE3	0.70	0.71			
	GE4	0.66	0.76			
	GE5	0.51	0.76			
EA	EA1	0.29	0.87	0.59	0.71	0.67
	EA2	0.37	0.91			
	EA3	0.21	0.66			

Table 3. Heterotrait-monotrait ratio of correlations

Variables	1	2	3
1. Motivation	--		
2. Game Engagement	0.93	--	
3. Environmental Attitudes	0.53	0.52	--

3 Discussion

The study aimed to validate an instrument designed to measure Motivation, Game Engagement, and Environmental Attitudes among younger students participating in digital game-based learning activities. The results from both the exploratory and confirmatory factor analyses provide support for the proposed three-factor structure, particularly for Motivation and Game Engagement, suggesting that the instrument is promising for educational contexts where digital games are used to support environmental and ocean literacy. The findings are consistent with prior research, which indicates that digital games foster meaningful learning through interactive features, immediate feedback, and engaging narratives [1, 2]. These characteristics are known to support exploration, decision-making, and emotional involvement during gameplay [3, 4]. The strong performance of the Motivation and Game Engagement scales reflects these dynamics: students typically experience gameplay as a stimulating and affectively rich activity, which may explain the high correlations between items within these dimensions. Additionally, the inclusion of environmental attitudes as a third dimension is supported by

literature showing that digital games can make abstract environmental concepts more concrete and accessible [5–7]. However, the psychometric performance of this scale was lower than that of the Motivation and Game Engagement dimensions. In particular, internal consistency was below desirable levels, and factor loadings were modest, suggesting that some items may require further semantic refinement for younger learners. Therefore, findings related to Environmental Attitudes should be interpreted with caution. The EFA results demonstrated a clear and interpretable factor structure, with most items loading on their intended factors, although the Environmental Attitudes items showed weaker performance. Sampling adequacy was ensured by the KMO values, and uniqueness values were within acceptable limits. These results are important because existing instruments, such as the Intrinsic Motivation Inventory [8–10], the Game Engagement Questionnaire [11], and environmental attitude scales [12, 13], were not originally designed for younger learners or for short, narrative-driven educational games. The present study addresses this gap by providing a measure that is age-appropriate, concise, and aligned with the cognitive and emotional demands of digital game-based learning experiences.

The CFA provided further confirmation of the model's validity. The fit indices indicated excellent model fit, and reliability values were stronger for Motivation and Game Engagement than for Environmental Attitudes. Convergent validity was supported for most factors, and although the discriminant validity between Motivation and Game Engagement was weaker, this outcome is theoretically consistent with digital game-based learning literature. Motivation and engagement often emerge simultaneously during gameplay and can be difficult to fully separate in self-reports by younger learners, who may tend to conceptualize both as part of a single, enjoyable experience [2]. This conceptual overlap does not undermine the model; rather, it reflects the nature of digital game-based learning contexts. The validated instrument offers a practical and theoretically grounded tool for researchers investigating how digital game-based activities shape students' motivational and affective responses, while the Environmental Attitudes dimension should be considered preliminary and in need of further refinement. As digital games become more deeply integrated into environmental education, the ability to measure these constructs reliably is essential. The instrument can support educators in evaluating the effectiveness of game-based interventions and adjusting their pedagogical strategies to better support students' emotional and cognitive involvement.

4 Limitations and Future Directions

While the psychometric results support for the proposed model, the study has some limitations. The high correlation between Motivation and Game Engagement suggests that the distinction between these constructs may be less clear for younger participants in digital game-based learning contexts. This may reflect the tendency of younger learners to experience motivational and engagement-related responses as part of a unified gameplay experience. Future research could explore whether more refined item word-

ing or alternative measurement approaches (e.g., observational data) can better differentiate these dimensions. A second limitation concerns the Environmental Attitudes subscale, which showed weaker psychometric performance than the other two dimensions, with lower internal consistency and more modest factor loadings. This suggests that some items may not yet capture the construct with sufficient clarity for younger students and may require further semantic refinement in future studies. In addition, the present study relied exclusively on self-report measures. Although this approach is appropriate for an initial psychometric validation, future research could complement these data with observational, cognitive, or behavioral indicators to provide a broader assessment of students' responses to digital game-based learning. Additionally, the instrument was validated within the context of environmental games. Future studies may test the scale in other content areas to examine whether the factor structure generalizes beyond environmental themes.

5 Conclusion

Overall, the study provides encouraging evidence that the newly developed instrument is suitable for assessing motivation, game engagement, and environmental attitudes among younger students engaged in digital game-based learning, while indicating the need for further refinement of the Environmental Attitudes subscale. Grounded in established theoretical frameworks [3, 8–10] and aligned with recent findings on the benefits of digital games for environmental education [5–7], this validated tool fills an important methodological gap and supports future research in the growing field of environmental digital game-based learning activities.

Acknowledgments. This work was carried out with the support of the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), funding code 88887.898520/2023-00, and is grounded in the methodology developed within the Literacy of the Contextual Use of Hyperconnection Technologies (LOCUST Project). It also received funding from national sources through FCT – Fundação para a Ciência e Tecnologia I.P., under the project/grant UID/50006 + LA/P/0094/2020 (doi.org/10.54499/LA/P/0094/2020), as well as from CECS via FCT (UIDP/00736/2020 + UIDB/00736/2020) through Portuguese national funds.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bado, N.: Game-based learning pedagogy: a review of the literature. *Interactive Learning Environments*. 30, 936–948 (2022). <https://doi.org/10.1080/10494820.2019.1683587>.
2. Plass, J.L., Homer, B.D., Kinzer, C.K.: Foundations of Game-Based Learning. *Educational Psychologist*. 50, 258–283 (2015). <https://doi.org/10.1080/00461520.2015.1122533>.

3. Gee, J.P.: What video games have to teach us about learning and literacy. *Comput. Entertain.* 1, 20–20 (2003). <https://doi.org/10.1145/950566.950595>.
4. Kiili, K., De Freitas, S., Arnab, S., Lainema, T.: The Design Principles for Flow Experience in Educational Games. *Procedia Computer Science.* 15, 78–91 (2012). <https://doi.org/10.1016/j.procs.2012.10.060>.
5. Koenigstein, S., Hentschel, L.-H., Heel, L.C., Drinkorn, C.: A game-based education approach for sustainable ocean development. *ICES Journal of Marine Science.* 77, 1629–1638 (2020). <https://doi.org/10.1093/icesjms/fsaa035>.
6. Leitão, R., Maguire, M., Turner, S., Arenas, F., Guimarães, L.: Ocean literacy gamified: A systematic evaluation of the effect of game elements on students' learning experience. *Environmental Education Research.* 28, 276–294 (2022). <https://doi.org/10.1080/13504622.2021.1986469>.
7. Tan, C.K.W., Nurul-Asna, H.: Serious games for environmental education. *Integrative Conservation.* 2, 19–42 (2023). <https://doi.org/10.1002/inc3.18>.
8. Deci, E.L., Ryan, R.M.: *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US, Boston, MA (1985). <https://doi.org/10.1007/978-1-4899-2271-7>.
9. Deci, E.L., Ryan, R.M.: The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry.* 11, 227–268 (2000). https://doi.org/10.1207/S15327965PLI1104_01.
10. Deci, E.L., Ryan, R.M.: Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology / Psychologie canadienne.* 49, 14–23 (2008). <https://doi.org/10.1037/0708-5591.49.1.14>.
11. Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., Pidruzny, J.N.: The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology.* 45, 624–634 (2009). <https://doi.org/10.1016/j.jesp.2009.02.016>.
12. Muñoz-García, I.M., Alcántara-Manzanares, J., Medina Quintana, S.: Key Aspects of Adolescents' Environmental Attitudes with a View to Transformative Education. *Education Sciences.* 12, 591 (2022). <https://doi.org/10.3390/educsci12090591>.
13. Wahyuningsih: The Attitude of Young People Towards Environmental Issues and Green Products. In: *Proceedings of the International Conference on Management, Accounting, and Economy (ICMAE 2020)*. Atlantis Press, Jakarta, Indonesia (2020). <https://doi.org/10.2991/aebmr.k.200915.033>.
14. Cheung, G.W., Rensvold, R.B.: Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal.* 9, 233–255 (2002). https://doi.org/10.1207/S15328007SEM0902_5.
15. Hu, L., Bentler, P.M.: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal.* 6, 1–55 (1999). <https://doi.org/10.1080/10705519909540118>.
16. Fornell, C., Larcker, D.F.: Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research.* 18, 39 (1981). <https://doi.org/10.2307/3151312>.
17. Henseler, J., Ringle, C.M., Sarstedt, M.: A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. of the Acad. Mark. Sci.* 43, 115–135 (2015). <https://doi.org/10.1007/s11747-014-0403-8>.

Digital to Physical Storytelling: From Digital Teacher Narratives to Tangible Student Engagement

Sofia Ribeiro¹[0000-0001-8665-1565], Pedro Beça¹[1111-2222-3333-4444] and Mónica Aresta¹

¹ DigiMedia – Digital Media and Interaction
University of Aveiro
Aveiro, Portugal
saribeiro@ua.pt

Abstract. Education increasingly uses hybrid methods, in which digital and physical elements complement each other to create meaningful learning experiences. Within this landscape, digital storytelling offers a pedagogical framework that structures learning through narrative sequences rather than through standalone multimedia artifacts, impacting on motivation and cognitive processes. This paper presents how the Dig2Phy platform integrates such narratives to support teachers in the design of educational resources, conceived to embed physical learning environments. Furthermore, it depicts a segment of a four-year design-based research, which unfolded in iterative development cycles with a participatory approach and exploratory nature, having the collaboration of in-service teachers from the 3rd Cycle of Basic Education and Secondary. The data that supports the evidence presented on this paper was collected through a national-wide survey with 237 valid responses, 14 semi-structured interviews, 5 usability tests, and 8 teachers that created educational resources in the Dig2Phy platform that were used, in the classroom, by 221 students. Findings suggest that the digital development model to promote digital competencies in the areas of Professional Engagement and Educational Resources (based on the DigCompEdu framework), embodied in the Dig2Phy platform, supports teachers in designing educational resources with compliance with design principles and pedagogical standards. Teachers' perceptions about students mentioned higher engagement and motivation towards learning and collaborative environments. Students described the tangible artifacts as more engaging and appealing, underscoring the added value of combining digital authoring with hands-on interaction. Overall, the research demonstrates that digital-to-physical storytelling is a powerful approach for integrating teacher digital development and student engagement.

Keywords. Digital storytelling, Digital authoring platform, Teachers' digital development model, DigCompEdu, Tangible educational resources.

1 Introduction

Learning is increasingly shaped by phygital ecosystems in which digital and physical elements complement each other to support meaningful student engagement [1], [2],

[3]. In contemporary educational contexts, digital storytelling has emerged as a powerful pedagogical strategy to enhance motivation, sustain attention and promote deeper cognitive involvement [4], [5], [6]. When narratives extend beyond the screen and take form as tangible artifacts that students can manipulate, analyze and connect with, storytelling becomes an embodied learning experience with strong emotional impact.

Simultaneously, as the educational landscape evolves, teachers increasingly assume the dual role of digital content creators and curators, facing barriers related to aesthetic standards, technical limitations, technological infrastructural deficiencies, time management challenges, and complexities concerning digital authorship and intellectual property rights [7], [8]. As digital technologies become inseparable from pedagogical practices [4], tools that support teachers in navigating these demands becomes urgent.

Addressing these challenges, this paper examines how a digital authoring platform – Dig2Phy – enables teachers to design digital learning sequences that materialize into physical educational resources to be used in the classroom or at students' homes. In this work, digital storytelling is not conceptualized as a multimedia artifact, but rather as a pedagogical structure that organizes learning through narrative sequencing. The Dig2Phy platform was developed to support both pre- and in-service teachers in creating tangible educational resources while simultaneously strengthening their digital competencies. Grounded in the DigCompEdu framework, particularly in the areas of Digital Resources and Professional Engagement, the Dig2Phy platform provides a flexible digital environment where story-driven learning pathways can be created and later be transformed into physical artifacts.

By involving teachers into the conception of this digital platform and analyze the perceptions of the students interacting with the tangible educational resources, this work contributes to understanding how digital narrative authoring platforms can empower teachers, as well as foster richer, more engaging classroom practices and higher student engagement. The following sections present the theoretical foundations of this approach, the research methodology, a description of the platform's features, the empirical results, and the implications for future development and research.

2 Background

This section examines how digital narratives and their underlying structures inform the design of meaningful learning materials. It also explores students' engagement with physical artefacts and the evolving role of teachers as digital authors or curators, establishing the conceptual basis for the Dig2Phy platform.

2.1 Digital learning sequences to enhance competency

Digital storytelling in educational contexts has evolved beyond multimedia artifacts, being also narrative learning sequences that structure tasks, decisions and reflective moments into a process that guides learners through a pedagogical rationale [9], [10], [11], [12]. In this paper, digital storytelling is approached as a digital narrative learning sequence rather than as a standalone multimedia artifact.

Learning sequences and narrative-centered learning shows that the order and structure of events impact immersion, motivation and cognitive processes, especially in collaborative and exploratory tasks [10]. Therefore, storytelling as learning sequences promotes motivation, engagement and active participation, being pivotal for educational tools [12], [13]. In fact, research points out that the order and structure of events significantly affect learning outcomes, emphasizing the pedagogical relevance of narrative sequencing [10].

2.2 Student Engagement through Physical Educational Resources

Physical educational resources play a significant role in sustaining attention, promoting collaboration and supporting hands-on activities. Studies involving tangible learning materials, such as LEGO Mindstorms and other construction-based activities, demonstrate improvements in student motivation, social interaction and problem-solving abilities [14], [15]. Tangible artifacts also support collaborative knowledge construction in activities, such as brainstorming, concept mapping and planning [16].

In contrast to the cognitive overload often associated with dense digital environments [17], [18], physical materials evoke emotional responses, stimulate debate and encourage in-person interaction [19], [20]. When digital information is embedded into tangible forms, phygital spaces emerge, blending the affordances of physical interaction with the flexibility of digital layers [21], [22]. These phygital environments are particularly effective in promoting collaboration, decision-making and long-term engagement [23], [24].

2.3 Teachers as Digital Authors and Curators

Despite growing access to digital technologies, effective fruitful use remains uneven [25], [26]. In fact, teachers frequently face technical (e.g., skills), pedagogical (e.g., educational potential), aesthetic (e.g., visual design knowledge), and authorial (e.g., content licensing) difficulties, as well constraints regarding the lack of technological infrastructure and technical support, often resulting in resources that lack visual consistency or pedagogical integrity [7], [8]. Authoring platforms can mitigate these barriers by providing intuitive interaction paradigms, embedded templates, open repositories, reusable educational resources and guided design suggestions [27], [28], [29], [30].

Furthermore, the rising production of user-generated educational resources intensifies questions surrounding copyright, licensing, and ethical reuse [31], [32]. Clear attribution mechanisms and integration of open licensing rights, such as Creative Commons, are paramount to support responsible copyrights [33]. As natural creators of their educational resources, teachers benefit from environments that simplify design processes with low levels of cognitive effort, enhanced interaction design, direct manipulation, customization and personalization, and open repositories with reusable educational resources, while aligning visual quality and pedagogical integrity [30].

The European Framework for the Digital Competence of Educators (DigCompEdu) provides a comprehensive reference for understanding and developing the digital com-

petencies in teachers require to effectively integrate technology into educational practices [34]. Based on this framework, the research addressed two areas, Professional Engagement and Digital Resources.

3 Methodology: Research process and studied sample

A design-based research methodology with an exploratory nature and participatory design approach was adopted, by involving integrated in-service teachers through each cycle to participate in the creation and validation of the digital development model. The methodological approach ensured alignment between the conceptual model – from previous research [35] and the literature review – and the practical constraints faced by teachers when creating educational resources – through surveys, interviews, usability tests, User eXperience methods, and validations in real classroom settings.

Throughout four years of research, 237 responses were collected in a national-wide survey, 14 semi-structured interviews were conducted, 5 usability tests were organized, and final validation in real context involved 221 students that used resources created in the Dig2Phy platform by 8 teachers. The in-service teachers who participated were from the 3rd Cycle of Basic Education and Secondary, from Portugal. Concerning ethical requirements of the data collection process, informed consent was delivered and signed, being authorized under GDPR regulation.

The national survey comprised four main sections: socioprofessional characterisation, pedagogical practices with digital platforms, digital competence self-assessment (adapted from the DigCompEdu framework), and identification of potential barriers to digital integration. Responses were collected on a Likert scale and categorised according to the six DigCompEdu proficiency levels (A1–C2). Semi-structured interviews explored teachers' pedagogical practices, existing resources, and receptiveness to the Dig2Phy concept. Usability tests employed the System Usability Scale (SUS) and the Self-Assessment Manikin (SAM) to evaluate perceived usability and affective response. In the real-context stage, teachers and students completed pre- and post-interaction surveys on digital competency and engagement, complemented by individual teacher interviews and a student feedback survey (applied by the teachers and reported back to the researcher).

4 The Dig2Phy Platform: From Digital to Physical

The Dig2Phy platform (see Fig. 1) acts as a narrative-driven authoring environment, enabling teachers to design digital educational resources that are later translated into physical artifacts for use in the classroom or at home by students.



Fig. 1. An overview of the Dig2Phy platform focusing on image direct manipulation, the homepage, and the color customization in the side menu, respectively.

Digital storytelling is conceptualized in the Dig2Phy platform as the step-by-step sequential logic that guides teachers in the creation of the educational resource: select layout-template, create texts, add images, and generate QRcode. During the research, this guided approach was emphasized [35] as pertinent for the creation of educational resources, reducing cognitive load as tasks are objective.

The educational resources can be completely tailored to specific class or students' needs since both the textual and media content is editable. As for design specifications, to ensure pedagogical integrity and compliance with design principles, users can select layouts from available templates and the color, which will be personalized into the resource by default standards, assuring contrast between text and background – based on the WCAG 2.0 standards, resorting to the WebAIM Contrast Checker API¹.

Ethical copyrighting is ensured since all images that can be used in the Dig2Phy are available within itself, having Shutterstock and Freepik as sources. Unfortunately, specific educational image libraries were not available to be integrated into third-party services, such as “Servier Medical Art”², “BioRender”³ or “Casa das Ciências”⁴. These would be extremely useful for visual disciplines, such as science-related disciplines or History. Furthermore, as the Dig2Phy platforms supports the remix of educational resources, it unequivocally identifies the original author of the educational resource as an automated process, without overburden users with that task.

As an all-in-one platform to support the creation of educational resources, teachers can even generate QRcode for external multimedia content, such as YouTube videos, all within the Dig2Phy platform. In the fourth step, users only have to paste the hyper-link of the content and the QRcode is generated and placed in the educational resource without covering previously added texts or images, aiming to promote users' autonomy and reduce cognitive load.

¹ <https://webaim.org/resources/contrastchecker/>, accessed in 2025/12/04

² <https://smart.servier.com/>, accessed in 2025/12/04

³ <https://www.biorender.com/>, accessed in 2025/12/04

⁴ <https://www.casadasciencias.org/>, accessed in 2025/12/04

Lastly, once finished the educational resource, a PDF file is produced with the resources prepared for printing and to be used as a physical artifact.

4.1 Narrative Learning Sequences to Promote Digital Competencies

As mentioned, the Dig2Phy platform guides the teacher in a logic sequence to create educational resources. Whilst avoiding cognitive overload, this narrative sequence promotes focus and deep learning – a process pivotal to the teachers' digital development model, since it aims to support teachers in fostering and enhancing digital competencies essential for their professional engagement.

Results from the real-context validation stage encompass multiple dimensions of assessment. Usability evaluation using the SUS scale yielded a score in the "Good" to "Excellent" range, and cognitive workload assessed via the NASA-RTLX registered a mean of 8 out of 20 ($SD = 1.94$), indicating moderate and consistent cognitive effort across participants – an expected level for a first-time creative authoring task. To evaluate the model's impact, teachers self-assessed their digital competencies with the DigCompEdu framework before and after interacting with the Dig2Phy platform. Table 1 summarizes the distribution across proficiency levels for the four DigCompEdu dimensions assessed.

Table 1. DigCompEdu proficiency levels before and after Dig2Phy interaction ($N = 8$).

DigCompEdu Dimension	Predominant Level Before	Predominant Level After	Key Shift
Professional Engagement: Organizational communication	B1–C1	B2–C1	Consolidation at B2
Professional Engagement: Professional Collaboration	A2–B1	B1–B2	Shift towards B1/B2
Digital Resources: Creating & Modifying	B1	B2	Upward movement to B2
Digital Resources: Managing, Protecting, Sharing	A2	B1–B2	Notable gain in sharing practices

The most notable shift was an evolution towards level B2 (Expert), which DigCompEdu identifies as the threshold for critical, innovative technology use [34]. Teachers specifically reported developing competencies in copyright-aware image sourcing and QR code creation they explicitly linked to the platform's embedded affordances. Six of the eight teachers felt that the platform supported digital competency development, with copyright awareness and QR code creation most frequently cited as new skills acquired.

4.2 Perceptions on Student Engagement

Drawing on the individual interviews conducted with all eight teachers after the classroom activity, teachers consistently described the tangible format as a differentiating

motivational element. Three teachers noted that the card-like format evoked a playful dimension students appreciated: one teacher reported students asking during breaks "Can we play cards?" (P1), while another observed that students actively used the resources as bookmarks in their textbooks after the activity (P6). Five teachers highlighted the platform's intuitiveness as the primary reason they would use it again – notably that the learning curve flattened significantly after the first resource was created (P2, P8).

Student feedback was collected from 206 of the 221 participants, as some were absent during the feedback session. On a 1–5 rating scale for overall activity quality, the majority rated it 4 ($n = 107$) or 5 ($n = 53$), with only 13 students giving a rating of 1 or 2. Regarding motivation, 137 students described the activity as "motivating" ($n = 94$) or "very motivating" ($n = 43$), while 46 found it "reasonably motivating". Among the positive aspects highlighted, students most frequently cited the activity's "fun" character ($n = 37$), knowledge acquisition and consolidation ($n = 31$), its innovative/differentiating nature ($n = 18$), and the clarity of information ($n = 12$). The QR code integration was noted positively by 7 students. The main negative aspects related to small font size ($n = 17$), print quality ($n = 7$), and QR code accessibility ($n = 6$) – all issues addressable through platform refinements in the printing stage.

5 Conclusions and Future Work

By enabling teachers to design pedagogically coherent resources that unfold physically in the classroom, the platform bridges design affordances and digital authoring with the motivational and collaborative benefits of tangible learning environments. This study shows that the platform effectively supported teachers in designing educational resources that can be materialized as tangible educational artifacts, fostering student engagement and knowledge acquisition. The Dig2Phy platform's features promote the development of teachers' digital competencies, essential for career progression and fruitful use of digital technologies – creating, curating and adapting educational resources, acquiring knowledge about intellectual property rights, while maintaining pedagogical integrity.

The results collected suggest that combining digital authoring processes with tangible educational resources may support both teachers and students. For teachers, the process encourages mainly the development of digital competencies related to the creation and adaptation of educational resources. For students, the physical format of the artefacts appears to enhance engagement and collaborative learning dynamics in the classroom. Furthermore, it also revealed increasing student motivation and interest towards the educational resources, a particularly relevant finding given the ongoing limitations in digital device availability and technological infrastructure in many Portuguese schools.

Future developments will focus on working with specific educational multimedia libraries to integrate more useful content for visual disciplines, AI-assisted content translation for remixing and tailored educational resources, and formal institutional

support to expand adoption of the Dig2Phy platform to measure its long-term educational impact.

Acknowledgments. This work was financially supported by national funds through FCT – Foundation for Science and Technology, I.P., under the project 2022.10021.BD.

References

1. Floyd, K. K., Shambaugh, N.: Applying Universal Design for Learning to Create a Transformational and Accessible Learning Framework for a Technology-Driven International University. In: Handbook of Research on Challenges and Opportunities in Launching a Technology-Driven International University, pp. 179–194 (2019). doi: 10.4018/978-1-5225-6255-9.ch010
2. ASLERD: Timisoara Declaration: Better Learning for a Better World through People Centred Smart Learning Ecosystems, http://www.mifav.uniroma2.it/in-event/events/aslerd/docs/TIMISOARA_DECLARATION_F.pdf, last accessed 2025/12/04
3. Fonseca, M. J., Mealha, Ó.: Student Smartphone Experience Narratives Mediated by the Phygital School Library for Learning Ecosystems. In: Dascalu, M., Marti, P., Pozzi, F. (eds.) Polyphonic Construction of Smart Learning Ecosystems. SLERD 2022. Smart Innovation, Systems and Technologies, vol. 908, pp. 13–28. Springer, Singapore (2023). doi: 10.1007/978-981-19-5240-1_2
4. Membrive, V., Armie, M., Membrive, V., Armie, M.: Story-Telling, Gamification, and Videogames: A Case Study to Teach English as a Second Language. In: Handbook of Research on Challenges and Opportunities in Launching a Technology-Driven International University, vol. 4–4, pp. 1436–1450. IGI Global (1AD). doi: 10.4018/978-1-6684-3710-0.CH069
5. Jiménez-Valverde, G.: Narrative Approaches in Science Education: From Conceptual Understanding to Applications in Chemistry and Gamification. Encyclopedia 5(3) (2025). doi: 10.3390/encyclopedia5030116
6. Campos, F. R. P.: Os professores como autores e editores de recursos educativos digitais: Uma investigação-ação na escola. Tese de Doutoramento, Universidade de Lisboa, Lisboa (2012), <http://hdl.handle.net/10451/8743>, last accessed 2025/12/04
7. Urquizo, J. et al.: Laying the Foundations for a Digital Literacy Program in the Galapagos Islands. In: 2020 IEEE Global Humanitarian Technology Conference (GHTC), pp. 1–8. IEEE (2020). doi: 10.1109/GHTC46280.2020.9342878
8. Rodríguez, C. L., García-Jiménez, M., Massó-Guijarro, B., Cruz-González, C.: Digital Storytelling in Education: A Systematic Review of the Literature. Rev Eur Stud 13(2), 13 (2021). doi: 10.5539/res.v13n2p13
9. Valle Torre, M., Oertel, C., Specht, M.: The Sequence Matters in Learning - A Systematic Literature Review. ACM International Conference Proceeding Series, 263–272 (2024). doi: 10.1145/3636555.3636880
10. Robin, B. R.: The Power of Digital Storytelling to Support Teaching and Learning. Digital Education Review 30, 17–29 (2016), <http://greav.ub.edu/der/>, last accessed 2025/12/04
11. Hung, C.-M., Hwang, G.-J., Huang, I.: A Project-based Digital Storytelling Approach for Improving Students' Learning Motivation, Problem-Solving Competence and Learning

- Achievement on JSTOR. *Educational Technology & Society* 15(4), 368–379 (2012), [suspicious link removed], last accessed 2025/12/04
12. Lippiello, S., Boscolo, A.: Adapting the DIST-M Model for Designing Experimental Activities—A Theoretical Discussion from an Interdisciplinary Perspective. *Education Sciences* 14(5), 472 (2024). doi: 10.3390/EDUCSCI14050472
 13. Afari, E., Swe Khine, M.: Robotics as an Educational Tool: Impact of Lego Mindstorms Article in *International Journal of Information and Education Technology*. *International Journal of Information and Education Technology* 7(5) (2017). doi: 10.18178/ijiet.2017.7.6.908
 14. Souza, I. M. L., Andrade, W. L., Sampaio, L. M. R., Araujo, A. L. S. O.: A systematic review on the use of LEGO® robotics in education. *Proceedings - Frontiers in Education Conference, FIE*, vol. 2018-October (2019). doi: 10.1109/FIE.2018.8658751
 15. Digiano, C., Tatar, D. G., Kireyev, K.: Learning from the Post-It: Building collective intelligence through lightweight, flexible technology (2006), https://www.researchgate.net/publication/228373384_Learning_from_the_Post-It_Building_collective_intelligence_through_lightweight_flexible_technology, last accessed 2023/06/08
 16. Aresta, M., Fernandes, A., Ribeiro, A. I.: Cenários de aprendizagem ubíqua: novos desafios. In: *Desafios Curriculares no Séc. XXI*, pp. 427–456. Coimbra (2022)
 17. Hawker, E. C.: Does Tangible design and print hold value in the contemporary digital age?. *Fields: journal of Huddersfield student research* 7(1) (2021). doi: 10.5920/fields.816
 18. Fuchsberger, V., Beuthel, J. M., Bentegeac, P., Tscheligi, M.: Grandparents and Grandchildren Meeting Online: The Role of Material Things in Remote Settings. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–14. ACM, New York, NY, USA (2021). doi: 10.1145/3411764.3445191
 19. Rafaeli, A., Vilnai-Yavetz, I.: Emotion as a Connection of Physical Artifacts and Organizations. *Organization Science* 15(6), 671–686 (2004). doi: 10.1287/ORSC.1040.0083
 20. Jung, H., Stolterman, E.: Material probe: Exploring materiality of digital artifacts. *Proceedings of the 5th International Conference on Tangible Embedded and Embodied Interaction, TEI'11*, 153–156 (2011). doi: 10.1145/1935701.1935731
 21. Costa, L. V., Veloso, A. I., Mealha, Ó.: A Review of Proxemics in ‘Smart Game-Playing’. In: *Smart Innovation, Systems and Technologies*, vol. 80, pp. 219–226. Springer Science and Business Media Deutschland GmbH (2018). doi: 10.1007/978-3-319-61322-2_22
 22. Silva, P., Moreira, I.: Curadoria em arte/arqueologia: processos de proto e pós-escavação. *Midas* 15 (2022). doi: 10.4000/midas.3574
 23. Torres, P. E. et al.: A systematic review of physical–digital play technology and developmentally relevant child behaviour. *Int J Child Comput Interact* 30, 100323 (2021). doi: 10.1016/j.ijcci.2021.100323
 24. Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt, E.: *Preparing for Life in a Digital Age*. Springer International Publishing, Cham (2014). doi: 10.1007/978-3-319-14222-7
 25. OECD: *OECD Skills Outlook 2021: Learning for life*. In: *OECD Skills Outlook*. OECD Publishing, Paris (2021). doi: 10.1787/0ae365b4-en
 26. Cross, N.: Expertise in design: an overview. *Des Stud* 25(5), 427–441 (2004). doi: 10.1016/J.DESTUD.2004.06.002

27. Pazmino, P. F. J., Slattery, B., Lyons, L., Hunt, B.: Designing for youth interpreter professional development: A sociotechnologically-framed participatory design approach. In: Proceedings of the 14th International Conference on Interaction Design and Children, pp. 1–10. ACM, New York, NY, USA (2015). doi: 10.1145/2771839.2771840
28. Shekoyan, V., Guo, W.: Moving towards problems assignments with reduced goal-specificity in introductory STEM courses. In: Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education, pp. 1–7. IEEE (2014). doi: 10.1109/ASEE-Zone1.2014.6820665
29. Ribeiro, S., Beça, P.: Digital authoring tools to create educational resources An analysis framework sustained by a benchmarking research. Iberian Conference on Information Systems and Technologies, CISTI, vol. 2022-June (2022). doi: 10.23919/CISTI54924.2022.9820377
30. Dwyer, T.: Media Convergence. Open University Press (2010). doi: 10.4324/9780203855485
31. Gordillo, A., Barra, E., Quemada, J.: Facilitating the creation of interactive multi-device Learning Objects using an online authoring tool. Proceedings - Frontiers in Education Conference, FIE, vol. 2015-February, no. February (2015). doi: 10.1109/FIE.2014.7044497
32. Bortoluzzi, M., Marenzi, I.: Web searches for learning: How language teachers search for online resources. *Lingue e Linguaggi* 23 (2017)
33. European Commission, Joint Research Centre, Redecker, C., Punie, Y.: European framework for the digital competence of educators – DigCompEdu. Publications Office (2017). doi: 10.2760/159770
34. Ribeiro, S.: A digital tool to create physical artifacts: The Gamers4Nature project. University of Aveiro (2021), <http://hdl.handle.net/10773/32542>, last accessed 2025/12/04



O *Storytelling* Digital como Estratégia de Cocriação e Envolvimento da Comunidade Escolar

Ana Velhinho¹ [0000-0001-9978-8317], Carminda Lomba² [0009-0002-1088-894X],
Sónia Samagaio² [0009-0004-0637-5134], Isabel Costa² [0009-0000-5889-050X], Teresa Marques²
[0009-0003-8167-1630], Pedro Almeida¹ [0000-0001-5878-3317] and Luís Pedro¹ [0000-0003-1763-8433]

¹ Digimedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal
{ana.velhinho, almeida, lpedro}@ua.pt

² Agrupamento de Escolas de Aveiro, Aveiro, Portugal
{carminda.lomba, sonia.samagaio,
isabel.costa, teresa.marques}@aeaveiro.pt

Resumo. Nas escolas João Afonso e Homem Cristo, em Aveiro, as Bibliotecas Escolares têm apostado na promoção das literacias digitais e na adoção de ferramentas tecnológicas para fomentar a colaboração e envolver a comunidade, destacando-se a criação de murais digitais sobre diversidade cultural, realizados por alunos de diferentes nacionalidades, e eventos como a Biblioteca Humana, dedicada à partilha intergeracional de experiências e histórias de vida. Outras estratégias desenvolvidas por professores, em contextos curriculares e extracurriculares, incluem o clube de jornalismo, oficinas de leitura e escrita, podcasts entre estudantes que personificam filósofos, vídeos de divulgação científica com apresentadores humanos e avatares, e rubricas de vídeos curtos para inspirar a cidadania ativa. O protocolo entre a Escola Homem Cristo e a Universidade de Aveiro visa potenciar estas dinâmicas através da plataforma Polariscopes, centrada no storytelling digital, mediante a criação colaborativa de histórias com conteúdos partilhados pelos utilizadores. O projeto “Homem Cristo” envolve a comunidade escolar e pretende reforçar a ligação com a cidade e o seu património, através da partilha de testemunhos multimédia. Este projeto inclui a documentação da vida do jornalista que dá nome à instituição, visando preservar a memória desta personalidade e do edifício histórico da escola, e constituirá um arquivo vivo de conteúdos digitais indexados, a ser utilizado como base para novas narrativas. Espera-se que este piloto contribua para consolidar a Polariscopes como recurso aberto a cidadãos e entidades, promovendo a cidadania, a cocriação digital e a preservação de memórias coletivas.

Palavras-chave: *Storytelling* digital, Recursos digitais, Cocriação, Cidadania digital, Memória Coletiva.

1 Introdução

A Polariscopes constitui uma plataforma de cocriação e visualização de memórias coletivas, desenvolvida pela Universidade de Aveiro em parceria com a Universidade Nova de Lisboa e a Câmara Municipal de Aveiro, no âmbito de um projeto de investigação e desenvolvimento. Destina-se a cidadãos, instituições e investigadores,

permitindo a partilha, exploração e construção de narrativas colaborativas sobre património cultural e eventos significativos, como efemérides, celebrações tradicionais e eventos culturais. O seu principal elemento diferenciador reside na promoção do *storytelling* colaborativo, possibilitando aos utilizadores a contribuição com imagens, vídeos, sons e textos, com vista ao enriquecimento de histórias existentes ou à criação de novas narrativas em projetos temáticos que configuram arquivos vivos. Ao agregar múltiplas perspetivas em galerias de conteúdos e histórias, a plataforma converte memórias individuais em experiências coletivas, oferecendo diversas abordagens temáticas e possibilidades de recombinação multimédia. Adicionalmente, visa fomentar a participação e ampliar a descoberta e partilha de memórias significativas que promovam a diversidade, a inclusão e a aprendizagem interativa, integrando mecanismos de visualização personalizada com base em preferências do utilizador [1], bem como recursos de Inteligência Artificial (IA) para recomendação de conteúdos e apoio à criação de novas narrativas [2].

No âmbito de um protocolo estabelecido entre a Universidade de Aveiro e o Agrupamento de Escolas de Aveiro, encontra-se em operacionalização o projeto escolar “Literacias e Património Local – Homem Cristo: jornalista e cidadão”, a desenvolver no ano letivo de 2025-2026 na Escola Secundária Homem Cristo. Este projeto visa, através da utilização da plataforma Polariscope, promover a literacia digital e mediática através de iniciativas de *Storytelling* Digital (SD). O desenho da iniciativa contempla a sua continuidade e expansão a diferentes níveis de ensino, disciplinas e anos letivos, enquanto componente do plano regular de atividades da Biblioteca Escolar, explorando a articulação entre Cultura, Formação e Parcerias com a Comunidade, no quadro da promoção da literacia digital.

O presente artigo encontra-se estruturado em cinco secções: a primeira apresenta os objetivos do estudo e procede à contextualização da plataforma Polariscope; a segunda aborda a missão das Bibliotecas Escolares na promoção da literacia digital; a terceira apresenta exemplos de dinâmicas de SD desenvolvidas na Escola Básica João Afonso e na Escola Secundária Homem Cristo; a quarta descreve a operacionalização do projeto escolar no qual será utilizada a plataforma Polariscope; e, por fim, as considerações finais sistematizam os principais fundamentos do estudo e perspetivas de desenvolvimentos futuros.

2 O papel das Bibliotecas Escolares na promoção da Literacia Digital

A literacia digital compreende um conjunto de conhecimentos, capacidades e atitudes que possibilitam a utilização crítica, informada, criativa e responsável das tecnologias digitais, articulando-se com a literacia mediática na interação ética com a informação e os media, em contextos online e offline [3]. No domínio educativo, este conceito ultrapassa o uso técnico de dispositivos e aplicações e integra competências de acesso, avaliação, produção e comunicação da informação, bem como a compreensão dos impactos sociais, éticos e culturais associados às tecnologias digitais [4].

No contexto escolar, a literacia digital desempenha um papel estruturante no desenvolvimento de competências essenciais para o século XXI, nomeadamente o pensamento crítico, a resolução de problemas, a comunicação e a colaboração. Estas competências encontram-se sistematizadas no Quadro Europeu de Competências Digitais para os Cidadãos (DigComp), que define áreas-chave como a literacia da informação e dos dados, a comunicação e colaboração digital, a criação de conteúdos, a segurança e a resolução de problemas [5]. A versão DigComp 3.0, publicada em 2025, atualiza este referencial ao integrar desafios emergentes, incluindo a Inteligência Artificial, a cibersegurança, os direitos digitais e a desinformação [6].

Em estreito alinhamento com este enquadramento, a atualização de 2025 do Manifesto da Biblioteca Escolar IFLA-UNESCO [7] reforça o papel das bibliotecas escolares na promoção da ética digital, da colaboração global, da sustentabilidade e da participação cívica. Para além do acesso a recursos informacionais, as bibliotecas escolares contribuem para o desenvolvimento de competências de pesquisa, avaliação crítica da informação, uso ético das fontes e produção de conhecimento em ambientes digitais. Em Portugal, esta missão é operacionalizada pela Rede de Bibliotecas Escolares, através do programa ProLiteracias [8], em articulação com o Plano Nacional de Literacia Mediática 2025–2029 [9].

Apesar de 29,9% da população portuguesa apresentar níveis de literacia digital acima do básico, valor superior à média da União Europeia, persistem desafios significativos no reforço destas competências [10]. Estudos centrados no papel das bibliotecas escolares na promoção da literacia mediática podem, deste modo, contribuir para evidenciar o seu papel enquanto espaços de literacias múltiplas, embora se identifiquem limitações no nível da formação, dos recursos e da integração curricular [11].

3 Dinâmicas de *Storytelling* Digital no Agrupamento de Escolas de Aveiro

A escolaridade obrigatória em Portugal pretende integrar a literacia digital de forma transversal ao currículo, promovendo práticas pedagógicas ativas e inclusivas. Neste contexto, a biblioteca escolar assume-se como um espaço estratégico de articulação entre escola e comunidade, funcionando como sala de aula permanente, orientada para o desenvolvimento de competências essenciais à cidadania digital. No Agrupamento de Escolas de Aveiro (AEA), as bibliotecas escolares têm reforçado este papel através de práticas criativas centradas nas literacias e na cidadania ativa, nomeadamente por meio de atividades de *storytelling* digital em contextos curriculares e extracurriculares.

O SD é uma prática multimodal que combina texto, imagem, áudio e vídeo na construção de narrativas com recurso a tecnologias digitais [12]. Na Educação, tem potencial para promover simultaneamente a aprendizagem de conteúdos e o desenvolvimento da literacia digital, integrando competências de escrita, oralidade e criatividade [13]. A criação de narrativas multimédia contribui para a formação de alunos enquanto leitores críticos e *prosumers* [14], isto é, indivíduos capazes de consumir e produzir conteúdos de forma consciente, em contextos que, na atualidade, são frequentemente marcados pela desinformação.

A revisão de literatura sobre o uso de SD na Educação reconhece-o como uma ferramenta pedagógica eficaz ao articular narrativa, tecnologia e reflexão crítica, alinhada com as competências do século XXI (DigComp) [15] [16]. Estes estudos evidenciam impactos positivos no envolvimento, na motivação e no desempenho académico, bem como no desenvolvimento de competências de comunicação, pensamento crítico e criatividade [15] [16]. Adicionalmente, o SD é valorizado enquanto estratégia de aprendizagem colaborativa e inclusiva que promove a literacia digital, bem como a literacia visual e artística [17], o pensamento histórico [18] e o desenvolvimento linguístico [19], nomeadamente em contextos multiculturais [20].

Neste enquadramento, apresentam-se, de seguida, exemplos de práticas de SD desenvolvidas pelas bibliotecas escolares da Escola Básica João Afonso e da Escola Secundária Homem Cristo, do AEA, que articulam cultura, formação e parcerias, promovendo o envolvimento da comunidade escolar e local.

Newsletter da Biblioteca. Enquanto instrumento de promoção da pesquisa, do pensamento crítico e da escrita, o jornalismo digital constitui um eixo central da divulgação das atividades do AEA. Para além do sítio web e das redes sociais institucionais, foi criada a Newsletter da Biblioteca¹, com recurso à ferramenta digital TRUE², disponibilizada pelo jornal Público para a gestão de jornais escolares. Esta iniciativa envolve alunos e docentes na relação crítica com fontes e factos, promovendo a consciencialização para a desinformação e a participação cívica. Os conteúdos são produzidos de forma colaborativa em secções temáticas como “Projeto em movimento”, “Leituras comVIDA”, “Ciência & Saber”, “Em Agenda” e “Vozes da Comunidade”. Estudos recentes analisam o programa Público na Escola, e os usos da plataforma TRUE, como um modelo relevante de literacia mediática, destacando o seu carácter transversal e multidimensional [21]. Para além das funcionalidades em uso, importa salientar que outras experiências estudadas com a plataforma TRUE [22] visaram ampliar o envolvimento dos estudantes e do agregado familiar, ao promover a interação com os jornais escolares em contexto doméstico através de vários dispositivos, como a televisão [23].

Biblioteca Humana. Alinhada com o Manifesto da Biblioteca Escolar [24], esta iniciativa extracurricular promove a biblioteca como espaço de identidade, memória e património local. Através de dinâmicas de partilha intergeracional, acolhe membros da comunidade, como escritores, artistas e famílias de diferentes culturas, valorizando experiências, saberes e tradições. A iniciativa reforça a abertura à comunidade e a visibilidade mútua entre contextos escolares e sociais.

Mural da Diversidade e Mapa das Nacionalidades. Perante a diversidade cultural do AEA, com 42 nacionalidades representadas, o Gabinete de Apoio ao Aluno Migrante (GAAMI), em articulação com a Biblioteca Escolar e a Mediadora Cultural, têm vindo

¹ <https://true.publico.pt/biblioteca-aeaveiro>

² TRUE – Trustworthy News and Related content for a Unified writing Environment, projeto de investigação financiado pelo COMPETE2020, resultante da parceria entre o Jornal Público, a Universidade de Aveiro e a MOG Technologies.

a desenvolver iniciativas de criação de conteúdos multimédia. Recorrendo a murais digitais desenvolvidos na ferramenta colaborativa Padlet, os alunos partilham tradições culturais do seu país de origem e apresentam-se através do Mapa das Nacionalidades³ e do Mural da Diversidade⁴. Estes contributos formam uma rede de histórias, línguas e patrimónios que promovem a integração e valorização da diversidade como património coletivo da comunidade escolar.

Divulgação Científica na Biblioteca. Desenvolvida no âmbito da disciplina de Físico-Química do 7.º ano da Escola Básica João Afonso, esta iniciativa consiste na produção de vídeos curtos de divulgação científica, publicados na secção “Ciência & Saber” da newsletter da Biblioteca. Os alunos assumem o papel de apresentadores, em articulação com apresentadores virtuais, explorando curiosidades e experiências científicas, ao mesmo tempo que desenvolvem competências de comunicação, literacia científica e digital. A participação é aberta à proposta de temas que podem ser apresentados por qualquer aluno. Como vídeo inaugural desta rubrica⁵, foi realizado um conteúdo de divulgação do espaço da biblioteca escolar, evidenciando a sua natureza polivalente.



Fig. 1. “Divulgação Científica na Biblioteca”, rubrica extracurricular desenvolvida em articulação com a disciplina de Físico-Química do 7º ano da Escola Básica João Afonso (Agrupamento de Escolas de Aveiro).

60 segundos. Esta iniciativa visa a produção de vídeos curtos sobre cidadania, mobilizando recursos técnicos e humanos do Clube de Ciência Tech Lab/LED da Escola Secundária Homem Cristo. Os alunos utilizam o estúdio audiovisual, com apoio de docentes e de pares, para desenvolver conteúdos multimédia, incluindo a seleção de atores entre alunos voluntários. Os Laboratórios de Educação Digital afirmam-se, neste contexto, como ambientes pedagógicos orientados para metodologias ativas, inovação curricular e desenvolvimento de competências digitais, colaboração e inclusão educativa [25].

Podcast filosófico. Desenvolvido no âmbito da disciplina de Filosofia do 12.º ano da Escola Secundária Homem Cristo, este trabalho curricular promove o pensamento crítico, a argumentação e a criatividade através do *storytelling* digital. A proposta evidenciou elevada motivação dos alunos e diversidade de abordagens narrativas, refletindo práticas contemporâneas de comunicação digital. Os resultados obtidos

³ <https://padlet.com/bibliotecahc/mapa-das-nacionalidades-1xgfvrbgi45711wy>

⁴ <https://padlet.com/bibliotecahc/mural-da-diversidade-juntos-constru-mos-uma-escola-mais-plur-rkyhabnvwl7r1xdn>

⁵ <https://youtu.be/sEwlfFnNTZY>

alinham-se com práticas de consumo e expressão individual dos alunos em *chats*, comunidades digitais e redes sociais, onde produzem fotos, vídeos, avatares, *stories* e *memes*, revelando uma apropriação ativa das tecnologias, não apenas como público, mas como criadores de conteúdos multimédia capazes de comunicar identidade, humor e informação. A proposta evidenciou ainda maior interesse pela temática em estudo, bem como elevada motivação para o trabalho colaborativo e investimento criativo na produção de conteúdos multimédia. O SD revelou-se um veículo flexível e eficaz que, apesar da premissa comum de exploração do diálogo filosófico entre Platão e Gettier sobre o problema do conhecimento, promoveu abordagens expressivas e narrativas diversificadas, contribuindo para o desenvolvimento de competências argumentativas e de pesquisa, a par da literacia digital e mediática.

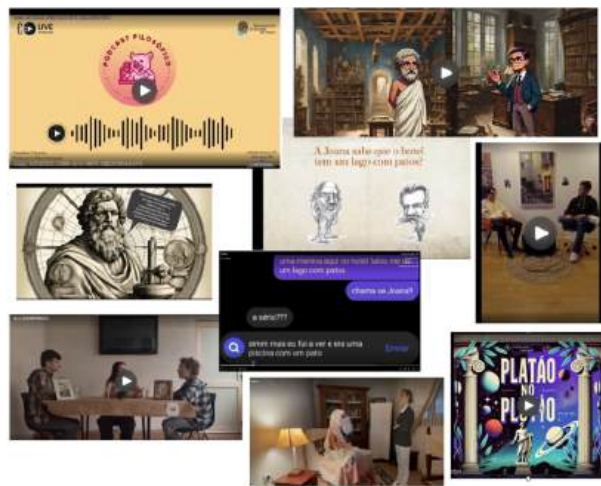


Fig. 2. “Podcast filosófico”, trabalho curricular desenvolvido no âmbito da disciplina de Filosofia do 12º ano da Escola Secundária Homem Cristo (Agrupamento de Escolas de Aveiro).

4 Operacionalização de projeto escolar em parceria com a plataforma Polariscope

Enquanto plataforma digital para a cocriação e visualização de memórias coletivas, a Polariscope pode ser utilizada em três contextos principais: pessoal, institucional e educativo. No contexto pessoal, indivíduos interessados em tradições, efemérides e experiências culturais podem explorar e partilhar conteúdos, promovendo a valorização da memória coletiva. No contexto institucional, equipas de museus, associações ou grupos culturais colaboram na organização e dinamização de eventos, integrando múltiplas perspetivas. O contexto educativo destaca-se pelo seu potencial pedagógico. Professores podem criar projetos colaborativos em que os alunos desenvolvem competências de investigação, análise crítica e trabalho em equipa, explorando conteúdos ligados ao património e a eventos culturais. A plataforma permite organizar conteúdos multimédia em projetos, criar histórias colaborativas, propor desafios pedagógicos e enriquecer as

narrativas com perspetivas diversificadas. Cada contribuição é protegida por Licença Creative Commons, assegurando crédito ao(s) autor(es) e uso responsável do material.

No âmbito da colaboração entre a Universidade de Aveiro e o AEA, o projeto escolar “Literacias e Património Local – Homem Cristo: jornalista e cidadão” envolve as 24 turmas dos três anos de escolaridade e cerca de 30 docentes, promovendo literacias, pensamento crítico, criatividade e participação ativa, funcionando também como prática de memória documental do patrono e do edifício histórico da escola. A integração da Polariscope nesta iniciativa tem como objetivos principais: i) homenagear o patrono da Escola Secundária Homem Cristo, valorizando a sua vida, obra e contributo cívico e jornalístico; ii) promover o conhecimento do património local e da memória histórica da escola, incluindo o edifício que a acolhe; iii) reforçar a identidade institucional e o sentido de pertença da comunidade educativa, através da documentação, preservação e divulgação destas memórias coletivas em acesso aberto, constituindo um arquivo digital dinâmico e em atualização contínua; iv) desenvolver competências de literacia informacional, mediática e digital, com enfoque no pensamento crítico, ético e na promoção da cidadania ativa; v) envolver a comunidade escolar e local, através da cooperação com bibliotecas, órgãos de comunicação social e entidades culturais, integrando metodologias de ensino baseadas em projetos, *storytelling* digital e cocriação de conteúdos em ambientes colaborativos.

4.1 Metodologia

A metodologia do projeto “Homem Cristo: jornalista e cidadão” assenta numa abordagem multidisciplinar, centrada no aluno e orientada pela biblioteca escolar, articulando práticas de literacia informacional e mediática com o currículo do Ensino Secundário. As atividades desenvolvem-se em contextos letivos e extraletivos, integrando disciplinas como Português, História, Filosofia, TIC e Cidadania e Desenvolvimento, e recorrendo aos espaços físicos e digitais da biblioteca como núcleo de investigação, produção e partilha.

A operacionalização organiza-se ao longo do ano letivo em sessões públicas com convidados, sessões de pesquisa orientada, análise crítica da informação, produção criativa e exposição dos resultados à comunidade. O trabalho decorre em pares e pequenos grupos, promovendo a colaboração, a responsabilidade partilhada e o desenvolvimento de competências sociais. A plataforma Polariscope assume um papel central enquanto ferramenta de curadoria, *storytelling* colaborativo e reflexão crítica sobre os media, permitindo a agregação de conteúdos multimédia, a construção de narrativas coletivas e a sua atualização contínua enquanto arquivo vivo e de acesso público, aberto a contributos externos à comunidade escolar e com continuidade para além do ano letivo.

Estão previstos momentos formativos específicos para alunos e docentes, através de oficinas e tertúlias sobre temas como verificação da informação, direitos de autor, ética digital e segurança online. Para além da plataforma Polariscope, a disseminação das atividades e recursos realiza-se através de uma exposição presencial, da newsletter da biblioteca escolar, do website e redes sociais oficiais do AEA, bem como da sua replicação pelos parceiros do projeto, incluindo a Biblioteca Municipal de Aveiro e

rádios e jornais locais. A monitorização incide nas dinâmicas de participação dos alunos, na qualidade dos processos, na integração curricular e na sustentabilidade das práticas, com vista a assegurar a continuidade e o impacto educativo.

4.2 Operacionalização e resultados esperados

A operacionalização do projeto organiza-se por ano de escolaridade e segundo eixos temáticos, em alinhamento com os objetivos de literacia, cidadania e produção crítica de conteúdos, culminando numa exposição final anual. Em articulação com os parceiros do projeto, procede-se, no final de cada ano letivo, à avaliação do impacto junto de alunos e docentes, com vista à adaptação das atividades e abordagens para o ciclo seguinte, bem como ao reforço e expansão das parcerias institucionais.

10.º ano | Vida e Contexto Histórico. Os alunos participam em sessões de literacia informacional, desenvolvendo competências de pesquisa, avaliação e validação de fontes. Produzem mapas conceptuais, linhas do tempo digitais e infografias de natureza biográfica e histórica, posteriormente divulgadas em murais digitais, na newsletter da biblioteca e na plataforma Polariscope.

11.º ano | Escrita Panfletária e Jornalismo. Privilegiam-se atividades centradas na escrita panfletária e no jornalismo, com ênfase na análise crítica dos media e no combate à desinformação. Os alunos elaboram artigos jornalísticos, podcasts e vídeos curtos sobre temas contemporâneos, refletindo sobre ética, direitos digitais e responsabilidade social. Estes conteúdos são objeto de curadoria e divulgação em redes sociais e meios locais, reforçando a dimensão pública do projeto.

12.º ano | Cidadania e Intervenção Social. As atividades incidem na cidadania e intervenção social, incluindo debates sobre liberdade de expressão e ética digital, bem como a produção de reportagens, documentários, instalações artísticas e exposições digitais. Esta etapa visa consolidar competências críticas, criativas e comunicacionais, promovendo a participação ativa e o envolvimento da comunidade escolar e local.

Enquanto produtos comunicacionais esperados, prevê-se: i) a integração de conteúdos multimédia em acesso aberto na plataforma Polariscope, constituindo um arquivo vivo com narrativas e curadorias temáticas abertas à participação externa; ii) a criação de painéis biográficos e linhas do tempo digitais sobre Homem Cristo; iii) o desenvolvimento de mapas conceptuais, infografias e murais digitais; iv) a produção de artigos jornalísticos e textos panfletários para a newsletter e plataformas digitais do agrupamento; v) a realização de podcasts com entrevistas, debates e reflexão sobre cidadania, media e ética da informação; vi) a produção de vídeos curtos, reportagens e documentários de natureza histórica, cívica e social; vii) a conceção de instalações artísticas, cartazes digitais e exposições com curadoria discente; viii) a criação de conteúdos para redes sociais; ix) a elaboração de um boletim jornalístico digital de continuidade do projeto; e x) a organização de uma exposição física aberta à comunidade, em parceria com a Biblioteca Municipal.

5 Considerações finais

Este caso de estudo evidencia o papel estratégico das bibliotecas escolares na promoção da literacia digital e mediática, através da integração de práticas de storytelling digital e de cocriação em contextos educativos formais e não formais. As experiências desenvolvidas no Agrupamento de Escolas de Aveiro demonstram que a articulação entre currículo, biblioteca escolar e comunidade potencia ambientes de aprendizagem participativos e orientados para a cidadania ativa. A diversidade de dinâmicas implementadas – incluindo jornalismo escolar, podcasts, vídeos de divulgação científica e iniciativas de partilha intergeracional – evidencia o potencial do storytelling digital como metodologia transversal para o desenvolvimento do pensamento crítico, da criatividade e da responsabilidade ética no uso da informação.

A utilização da plataforma Polariscope no projeto “Homem Cristo: jornalista e cidadão” constitui um contributo inovador, ao permitir a agregação, curadoria e recombinação de conteúdos multimédia produzidos colaborativamente, transformando memórias individuais em narrativas coletivas e em arquivos vivos de património local abertos à participação. Este piloto reforça a pertinência da plataforma enquanto recurso de acesso aberto a escolas, cidadãos e instituições culturais, promovendo a preservação da memória, a participação cívica e a aprendizagem ativa. Os resultados esperados apontam para a sustentabilidade e escalabilidade do modelo, sublinhando a relevância de parcerias entre escolas, ensino superior e comunidade na resposta aos desafios da educação digital e da literacia dos media.

Como desafio futuro, destaca-se a necessidade de reforçar a biblioteca escolar como espaço de inovação, participação e cidadania ativa, capaz de responder às exigências do ecossistema digital contemporâneo. Neste âmbito, importa refletir sobre a mobilização ética e inclusiva de narrativas digitais e de inteligência artificial, de modo a dar voz às comunidades escolares, valorizar a diversidade de experiências e transformar memórias individuais em conhecimento coletivo, reforçando a participação cívica, crítica e criativa.

Agradecimentos. Este trabalho é financiado por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto 2022.04424.PTDC e da UIDB/05460/2020.

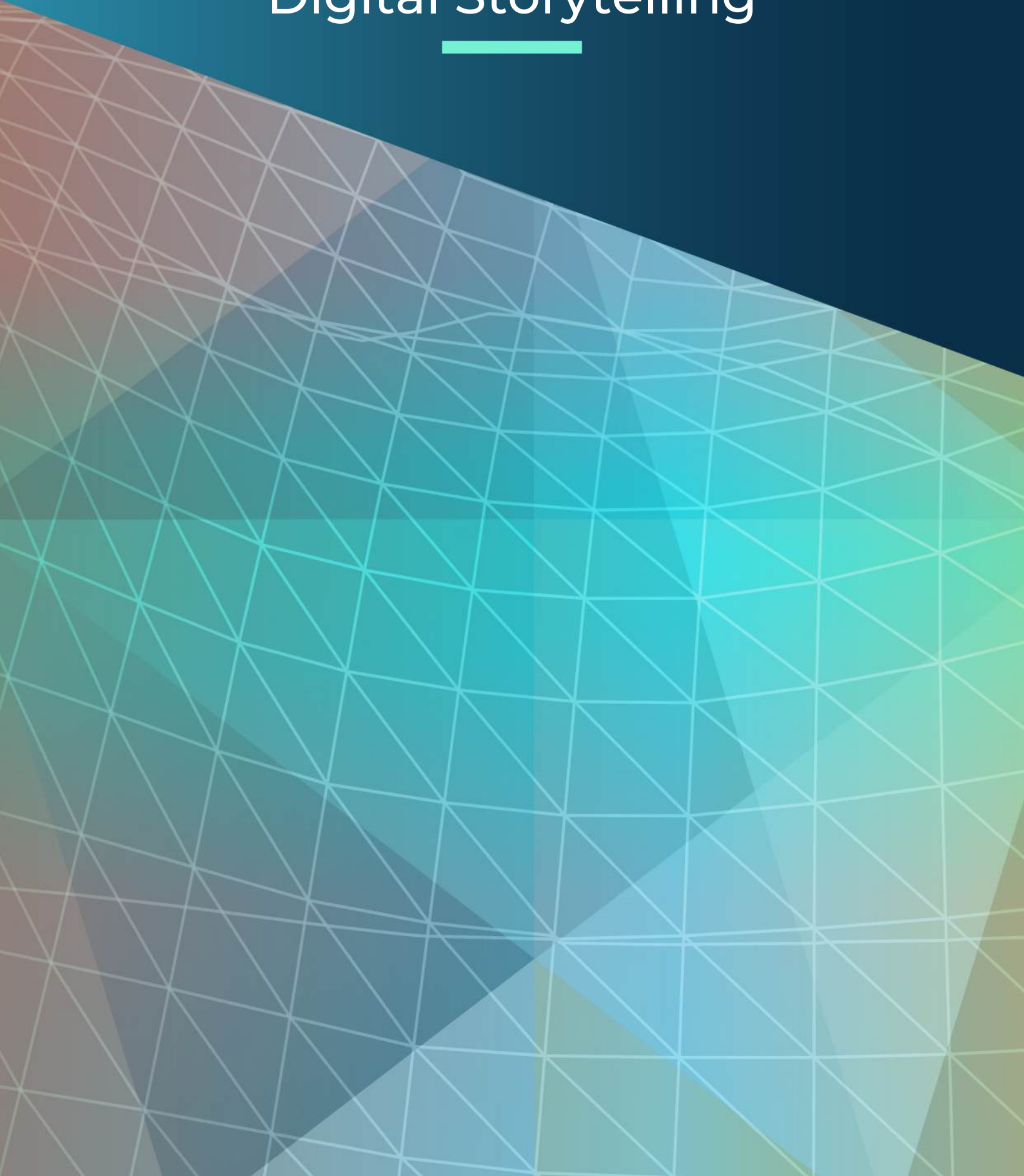
Referências

1. Roque, J., Almeida, P., Silva, T., Velhinho, A., Alves, M.: Personalized story-viewing experiences on the Polariscope platform. In: *Proc. ACM Int. Conf. Interactive Media Experiences Workshops (IMXw 2025)*, pp. 37–41. SBC, Porto Alegre (2025). <https://doi.org/10.5753/imxw.2025.3888>
2. Almeida, P., Teixeira, A., Velhinho, A., Raposo, R., Silva, T., Pedro, L.: Remixing and repurposing cultural heritage archives through a collaborative and AI-generated storytelling digital platform. In: *Proc. ACM Int. Conf. Interactive Media Experiences Workshops (IMXw 2024)*, pp. 100–104. ACM, New York (2024). <https://doi.org/10.1145/3672406.3672419>

3. UNESCO: Media and Information Literacy (MIL) – Facts and figures. UNESCO (n.d.). <https://www.unesco.org/en/media-information-literacy>
4. UNESCO: *A global framework of reference on digital literacy skills for indicator 4.4.2*. UNESCO (2018). <https://unesdoc.unesco.org/ark:/48223/pf0000265403>
5. European Commission, Joint Research Centre: The Digital Competence Framework (DigComp). European Commission (n.d.). https://joint-research-centre.ec.europa.eu/digcomp_en
6. European Commission, Joint Research Centre: *DigComp 3.0: European Digital Competence Framework*. Publications Office of the European Union (2025). <https://publications.jrc.ec.europa.eu/repository/handle/JRC144121>
7. International Federation of Library Associations and Institutions: *Manifesto da Biblioteca Escolar da IFLA-UNESCO 2025*. IFLA (2025). <https://repository.ifla.org/handle/20.500.14598/4026>
8. RBE/MEC: *Enquadramento – ProLiteracias: Media e Informação com a Biblioteca Escolar*. Rede de Bibliotecas Escolares (2025a). <https://www.rbe.mec.pt/np4/proliteracias-enquadramento.html>
9. RBE/MEC: *Convocatória às bibliotecas – Plano Nacional de Literacia Mediática 2025–2029*. Rede de Bibliotecas Escolares (2025b). <https://www.rbe.mec.pt/np4/lidera-enquadramento.html>
10. ANACOM: *Competências digitais da população e das empresas*. Autoridade Nacional de Comunicações (2024). https://anacom.pt/streaming/Competencias_digitais_populacao_e_empresas_2024.pdf
11. Pereira, S., Toscano, M.: *Literacia mediática nas bibliotecas escolares*. Observatório de Media, Informação e Literacia (MILObs); Rede de Bibliotecas Escolares (2021). <https://mi-lobs.pt/wp-content/uploads/2021/03/Literacia-Mediática-nas-Bibliotecas-Escolares-Estudo-MILObs-RBE.pdf>
12. Lambert, J.: *Digital storytelling: Cookbook*. Center for Digital Storytelling, Berkeley, CA (2010).
13. Hack, J.R., Guedes, O.: Digital storytelling, educação superior e literacia digital. *Roteiro* **38**(1), 9–31 (2013).
14. Toffler, A.: *The Third Wave*. Bantam Books, New York (1981).
15. Rodríguez, C.L., García Jiménez, M., Massó Guijarro, B., Cruz González, C.: Digital storytelling in education: A systematic review of the literature. *Review of European Studies* **13**(2), 13–26 (2021). <https://doi.org/10.5539/res.v13n2p13>
16. Nair, V., Yunus, M.M.: A systematic review of digital storytelling in improving speaking skills. *Sustainability* **13**(17), 9829 (2021). <https://doi.org/10.3390/su13179829>
17. Petousi, D., Katifori, A., Servi, K., Roussou, M., Ioannidis, Y.: History education done different: A collaborative interactive digital storytelling approach for remote learners. *Frontiers in Education* **12**, 1–19 (2022).
18. Castro, J.C.R.: “*Narro, logo existo!*”: *O storytelling e o digital storytelling como estratégias de desenvolvimento da literacia histórica dos alunos*. Master’s thesis, Faculdade de Letras, Universidade do Porto (2024). <https://repositorio-aberto.up.pt/bitstream/10216/161608/2/686788.pdf>
19. Ferreira, L.: *As vantagens do digital storytelling sobre a forma tradicional de contar histórias em contexto de ensino de Inglês no 1.º Ciclo do Ensino Básico*. Internship report, Escola Superior de Educação do Instituto Politécnico do Porto (2022).
20. Manganello, F., Baldacci, M.: Digital stories and inclusive cultures at school: A research study in an Italian primary multicultural classroom. *Education Sciences* **14**(10), 1108 (2024). <https://doi.org/10.3390/educsci14101108>

21. Reis, C.M.: “*Pequenas revoluções*”: *O papel dos órgãos de comunicação social na promoção da literacia mediática e educação para os media: o caso do Público na Escola*. Master’s thesis, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa (2024). https://run.unl.pt/bitstream/10362/179018/1/Mestrado_Cristiana_Matos_Reis_vcm.pdf
22. Júlio, L., Carvalho, D.M., Ulisses, A., Lima, I.R., Silva, T.: Engage students in news writing. *Procedia Computer Science* **219**, 999–1005 (2023). <https://doi.org/10.1016/j.procs.2023.01.377>
23. Bentes, S., Júlio, L., Velhinho, A., Encarnação, J., Mota, M., Carvalho, D.M., Santos, R., Silva, T.: TV notifications to promote trustworthy news and local school newspapers. In: *Lecture Notes in Networks and Systems*, pp. 242–248. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45642-8_24
24. IFLA, UNESCO: *Manifesto da Biblioteca Pública IFLA-UNESCO*. Direção-Geral do Livro, dos Arquivos e das Bibliotecas (2022). <http://bibliotecas.dglab.gov.pt/pt/etc/Paginas/ManifestoUNESCOsobreBibliotecasPublicas.aspx>
25. Direção-Geral da Educação: Laboratórios de Educação Digital. DGE (n.d.). <https://digital.dge.mec.pt/laboratorios-de-educacao-digital>

Applications for Digital Storytelling





A Framework for AI-Assisted Storytelling from Collective Multimedia Records

Rodrigo Graça^[0009-0003-1208-0642], Pedro Almeida^[0000-0001-5878-3317], Telmo Silva^[0000-0001-9383-7659], Rui Raposo^[0000-0002-5275-6200], Ana Velhinho^[0000-0001-9978-8317]
and Ana Teixeira^[0000-0003-0313-6408]

Digimedia, Department of Communication and Art, University of Aveiro, Aveiro, Portugal
{rodrigograca, almt00, almeida, tsilva, raposo, ana.velhinho}@ua.pt

Abstract. Digital Storytelling (DS) is an evolving and fundamental form of narrative that is crucial for cultural expression and collaborative content creation, particularly on web-based platforms. Despite the potential of modern co-creation environments, such as the Polariscope platform, users often struggle with developing full, high-quality narratives due to constraints in time, patience, or creative confidence. Addressing this critical barrier, this research explores the integration of Generative Artificial Intelligence (GenAI) as an assistive technology to optimize and facilitate the creative workflow. This research focuses on integrating GenAI into the Polariscope platform, aiming to enrich the creation and co-creation of stories, as well as content metadata and recommendations. It is intended to develop and analyze an intelligent assistant capable of supporting users in writing, generating descriptions for content, and automatically creating stories, utilizing the platform's existing content. Overall, the study aims to investigate the extent to which Artificial Intelligence (AI) can enhance the creativity, efficiency, and quality of generated stories, as well as to assess user interest and engagement with this type of creative assistance.

Keywords: Artificial Intelligence, Co-creation, Digital Storytelling, Creativity, Stories, Intelligent Assistant.

1 Introduction

The Polariscope is a Research and Development (R&D) project focused on establishing a collaborative digital storytelling platform. Its primary goal is to bring together citizens and cultural heritage institutions to share archives and memories, thereby promoting their co-creation and visualization. To enhance this process, the project proposes an AI-assisted solution capable of supporting users in the creation of narratives and automatically generating stories based on existing content. In the platform's context, a story is part of the data aggregation and organization hierarchy, which is structured into three levels: Projects, Stories, and Content. The highest level is the Project (e.g., a collective cultural event, such as a national Independence Day celebration). Within each Project we have the Stories, which are the narratives designed by the creators, encompassing

text, content, and memories; this is the level where the AI assistant is designed to provide support. At the base are the Contents, which include photos, videos, audio files, or citations that users share, and which can later be used within the Stories.

In this context, a critical challenge is the creation of Stories. Based on previous testing [1][2], it has been observed that most users tend to prefer either posting content or simply viewing what others have published. This behavior imposes significant limitations for the use of Polariscope, as the story creation is a longer process demanding more attention, time, and creativity. This possible limitation establishes the basis for the current research. Given the significant evolution of GenAI in recent years and the ongoing emergence of various tools supporting digital storytelling, there is a clear opportunity to use GenAI to assist creators in the co-creation of stories, recommending content, aiding in narrative writing, or even generating stories automatically.

The study is guided by the core research question: “Can the GenAI assistant significantly contribute to the story creation process and to its perceived quality?” This is encapsulated in the general objective of the research of developing a tool for the Polariscope platform to ease and improve the process of creating stories, through the integration of AI. This research adopts the Design-Based Research (DBR) methodology. This means the study is focused not merely on technological development, but on the iterative design and practical improvement of the assistant, emphasizing the user experience to ensure the process is made as easy and intuitive as possible for the story creators.

This paper is structured into six sections: the first establishes the research context; the second presents insights from previous work; the third provides an overview of the state of the art; the fourth details the methodology and its phases; the fifth outlines the proposed solution; and the final section provides a summary of the study and directions for future work.

2 Previous Work

Key findings from prior research [1][2] related to AI-assisted storytelling conducted within the Polariscope R&D project established relevant insights. In this preliminary research, the team aimed to find the optimal value of the AI’s Temperature parameter, which governs its level of creativity. The findings established that this parameter cannot be fixed. In these tests employing the GPT 3.5 model, users demonstrated varying preferences that were directly linked to the AI’s output creativity. Some participants favored stories generated with a high Temperature, indicating a preference for more creative and engaging narratives. Conversely, other participants preferred stories generated with a low Temperature, resulting in more formal and objective texts. We hypothesize that this divergence is driven by the fact that different narrative contexts demand distinct levels of factual adherence. For instance, a story about a local festive event, which often appeals to personal sentiment, showed a clear preference for the higher-temperature texts. Conversely, a narrative concerning a historical event, which inherently requires rigor and objectivity, led users to prefer the lower-temperature text.

This discovery directly informs the design of the AI assistant and poses the fundamental challenge that this research must address: how will the system dynamically decide the appropriate level of AI creativity within the context of the Polariscope platform? This issue is critical, as a one-size-fits-all approach to generative output would fail to meet user expectations across the varied story types within the platform's archives. This challenge directly shapes the requirements and evaluation metrics for the intelligent assistant proposed in this paper.

3 State of the Art

This work aims to develop an intelligent tool within the Polariscope platform to ease and enhance the story creation process through the integration of AI. To support this objective and the conceptual and theoretical foundation a literature review was prepared around three core pillars:

- **Digital Storytelling.** Establishes the creative context by defining DS, the concept of collaboration, and the operational role of co-creation platforms.
- **Generative AI.** Explores the technological domain, addressing the concept, evolution, and distinction of GenAI from traditional AI, along with its specific applications in creative and educational fields.
- **AI Assistant in Co-creation.** Bridges the two domains, focusing on Human-AI co-creation in narrative generation and analyzing the potential of AI assistants as optimizing tools within the creative process.

Digital Storytelling (DS), also referred to as Digital Narrative, has emerged as a significant evolution of traditional storytelling practices by integrating multimedia elements such as text, image, audio, video, and animation within digital environments. Rooted in the long-standing tradition of oral narration, DS bridges narrative heritage with contemporary digital affordances, enabling more immersive, participatory, and interactive forms of meaning-making. Unlike linear, printed storytelling, digital storytelling is inherently interactive and web-oriented, allowing audiences to engage through comments, annotations, and other participatory mechanisms embedded within the narrative itself [3]. As a transdisciplinary field, DS increasingly incorporates computational methods and emerging technologies, positioning itself at the intersection of narrative theory, media studies, human-computer interaction, and creative practice [4].

Collaboration constitutes a foundational principle within digital storytelling ecosystems. Across disciplines, collaboration is understood as a process in which multiple agents interact continuously, share resources, and jointly construct outcomes that exceed the potential of individual contributions [5][6]. In creative contexts, collaboration functions not merely as a means of efficiency but as a driver of innovation and experimentation. This collaborative logic is operationalized through co-creation platforms, digital systems designed to facilitate collective creativity and shared authorship [7]. However, prior research highlights persistent challenges in designing platforms that effectively support collaborative creative work, including coordination difficulties, un-

even participation, and cognitive overload [8]. As a result, many co-creation environments require facilitation mechanisms, increasingly envisioned as digital partners, such as AI assistants that can guide users, structure workflows, and enhance creative efficiency.

In parallel, Generative Artificial Intelligence (GenAI) has rapidly transformed the landscape of digital content creation. Distinguished by its ability to generate novel content that resembles human-produced data, GenAI represents a qualitative shift from earlier AI paradigms [9][10]. The field has evolved from early generative models such as Generative Adversarial Networks (GANs) to large-scale multimodal systems, including large language models like ChatGPT, capable of producing coherent text, images, and video from minimal input [9]. Unlike traditional AI and machine learning approaches, which are predominantly discriminative and focused on classification or prediction tasks, GenAI is characterized by its capacity to synthesize new and original outputs by learning underlying data structures rather than merely interpreting existing data [11][12]. This creative capability positions GenAI as a critical assistive technology for storytelling and co-creation contexts.

The impact of GenAI is particularly visible in creative and educational domains, where it introduces new opportunities for personalization, ideation, and process optimization [13][14]. In educational settings, GenAI can enrich learning experiences and support teaching practices through content generation and adaptive feedback [15]. In creative workflows, its role extends beyond automation, enabling creators to explore alternative ideas, accelerate production, and experiment with narrative forms. This evolution has reframed GenAI systems as assistants or collaborative partners rather than replacements for human creators [16]. Empirical studies of content creators, for instance on platforms such as YouTube, show that GenAI is commonly used to support specific tasks, including research, summarization, script drafting, and idea generation, while creative control remains firmly with the user [17].

At the intersection of Digital Storytelling, collaboration, and GenAI lies the paradigm of Human-AI co-creation. In this model, AI systems function as active partners that complement human creativity through suggestion, scaffolding, and rapid content generation [18]. The value of this collaboration lies in its capacity to combine computational efficiency with human intentionality, judgment, and cultural sensitivity. GenAI can introduce unexpected ideas, prompt creative exploration, and support problem-solving, while users retain agency by selecting, refining, or rejecting AI-generated inputs [19]. This synergistic relationship has become central to contemporary co-creation platforms, which increasingly integrate AI-driven functionalities such as draft generation, summarization, and narrative refinement [16][20].

Despite significant advances, a review of existing AI-powered storytelling tools reveals notable gaps. Fiction-oriented tools such as Sudowrite, AI Dungeon, NovelAI, and Artflow excel in imaginative content generation but lack robust support for uploading, contextualizing, and remixing large volumes of user-generated or archival material. Conversely, rigor-focused tools such as Grammarly or Canva Magic Write prioritize polishing, factual coherence, and multimodal integration, yet do not fully support collaborative storytelling grounded in archival or documentary content. This benchmark-

ing highlights the absence of integrated solutions that combine collaborative digital storytelling, GenAI-assisted co-creation, and the enrichment of existing multimedia archives. Addressing this gap represents a key opportunity for advancing the state of the art in AI-supported storytelling platforms and confirms the relevance of developing tailored AI assistants that balance creative support with user control and contextual rigor.

Ultimately, the study seeks to ideate and develop an intelligent assistant capable of generating narratives from existing content. The core goal is to evaluate the quality and user satisfaction of this support from the creators' perspective, while ensuring that users maintain creative control over the final output.

3.1 Digital Storytelling

Digital Storytelling also known as Digital Narrative, is an evolving narrative form that combines multimedia tools, such as images, videos, and animations, to create narratives that are more engaging and interactive than conventional ones. Inspired by the tradition of oral narration, DS has evolved with the advance of digital technologies, consolidating itself as an innovative practice that connects the narrative past to the expressive possibilities of the modern era.

Defining Digital Storytelling. Digital Storytelling refers to the use of media types, including voice, image, video, music, and graphics, in the construction of narratives, namely developed for web-based contexts [3]. Differentiating itself from traditional printed storytelling, DS promotes a more direct and engaging interaction with the audience, who can participate through comments, annotations and other interactive resources, often incorporated into the digital element itself [3]. It is a transdisciplinary field of research, characterized by a growing integration of computational methods and emerging technologies in the creation of digital narratives [4].

The Concept of Collaboration. Collaboration is a structural concept in various domains of study, characterized as the essential process where multiple agents work together to achieve a common goal [5]. The value of this practice lies in its ability to potentiate results that would be unattainable through isolated efforts [6]. In the creative context, collaboration is not just an efficiency tool, but a catalyst for innovation. In summary, collaboration is defined by continuous interaction, resource sharing, and the joint construction of a final product.

Co-creation Platforms. Co-creation platforms represent a digital mechanism that uses the concept of collaboration, allowing interaction and joint work among users [7]. However, the design of systems that support collaborative creativity faces significant challenges [8]. The implementation of these platforms often requires the presence of external facilitators or support mechanisms that guide the creative process and minimize the

challenges that are common in group work. This requirement for a support agent is fundamental for this study, as it establishes the need for a partner, namely a digital one, for example an AI assistant, which can promote efficiency and guide the user in content creation.

3.2 Generative AI

GenAI is an area of AI that is distinguished by its capacity to create new content that resembles human-generated data [9]. Its rapid development in recent decades has marked the beginning of a new era in digital content production, driving rapid growth in AI-generated content [10].

The field has progressed significantly from the initial development of models, such as Generative Adversarial Networks (GANs), to the emergence of large-scale models, including ChatGPT [9]. The most notable turning point of GenAI lies in its capacity to generate complex and varied content, which can include text, image, and video, from simple inputs [9]. This capacity of generative AI to create complex and coherent content at an unprecedented rate is what positions it as a crucial assistive technology.

Defining Digital Storytelling. Distinction from Traditional AI. The distinction between traditional AI and generative AI is crucial for understanding their application in co-creation. Traditional AI and Machine Learning (ML) approaches are often focused on discriminative tasks, where the primary objective is to analyze, classify, or predict based on existing data [11]. In contrast, GenAI differentiates itself by having the capacity to synthesize new and original content that did not previously exist in the training dataset [12]. The objective of GenAI is not just to interpret data, but to learn its underlying structures to generate unique outputs. This capacity for creation, rather than mere classification or prediction, makes GenAI the ideal technology for the project.

Creative and Educational Applications. The rapid evolution of generative AI has led to its widespread adoption across various sectors, with notable implications in the creative and educational fields [13]. GenAI presents significant opportunities and challenges for education, requiring schools to develop strategies for its effective integration [14]. The capacity of GenAI to produce text, images, and other types of media enables the enrichment of the learning experience and provides additional support to the teaching process [15]. Regarding creative applications, GenAI goes beyond mere efficiency. GenAI is a transformative technology that offers new opportunities for personalization, creation, and optimization in creative and educational environments [14].

AI Assistants in Human-AI Co-creation. The evolution of GenAI has transformed its role from a simple algorithm to that of an assistant or direct partner in the content creation process [16]. This perspective positions GenAI not as a substitute but as a support tool that optimizes tasks and increases the user's creative efficiency. The primary value of these assistants lies in their ability to generate and manipulate content rapidly, which is particularly relevant in the domain of User-Generated Content [16]. Research on con-

tent creators on platforms like YouTube demonstrates that users apply GenAI to optimize specific aspects of their workflow, such as script research, title generation, summarization, and idea inspiration [17].

Technical Context and Model Benchmarking. Considering the preliminary findings [1][2] and the goal of developing an optimized assistant for the Polariscope platform, a robust technical benchmarking was conducted to determine the best AI implementation approach. Following prior research, the solution involved adding two crucial parameters for output control: Top-p, which restricts the generated vocabulary, focusing the result on the intended theme, and Repetition Penalty, which forces the model to vary the words used within the chosen theme. Large Language Models (LLMs) such as ChatGPT 4-o and Gemini were identified as the main candidates, due to their proven excellence and their robust and well-known APIs. While open models like Llama 3 were also considered, the necessary hardware cost (especially the RAM on the virtual machine) would negate any expected cost benefits.

3.3 AI in cocreation

After establishing the conceptual pillars of Digital Storytelling and Collaboration (Section 3.1) and framing the technological advancement of Generative AI (Section 3.2), this section focuses on their intersection. The objective is to analyze how AI models can be integrated into digital platforms, acting as an active assistant and partner in the story creation process. The emphasis is placed on Human-AI co-creation models and how AI can optimize the creative workflow while simultaneously ensuring user control over the final result.

Human-AI Co-creation. This type of co-creation represents a collaboration paradigm where Generative AI systems evolve beyond a simple tool, taking on the role of an active partner in the creative process [18]. This collaboration is considered the most effective way to integrate technology into creation, generating synergies that combine the machine's processing capacity with human intention and sensibility [18]. The fundamental value of GenAI in this context lies in its capacity to stimulate creativity and problem-solving. AI can suggest ideas that would otherwise not occur to users, allowing the creator to select, refine, or use these inputs to generate new solutions [19]. This intervention is viewed as a scaffolding that guides users through a series of creative stages, enhancing their productivity and the novelty of their ideas [19]. In summary, Human-AI co-creation is a model that maximizes creative potential through the combination of AI's idea generation capabilities and the user's creative control [19].

AI-powered Co-creation Platforms. Co-creation platforms have evolved significantly with the integration of GenAI, transforming into ecosystems that unite the machine's generation capacity with human creative control [16][20]. The focus is on utilizing AI to optimize collaborative creation through functionalities such as draft generation, input summarization, and improvement of narrative quality. The importance of evaluating these tools is crucial, requiring a comparative analysis of different models to determine

their effectiveness in terms of functionality and ease of use in collaboration [20]. Thus, the success of digital co-creation critically depends on the strategic choice and integration of an AI assistant that can facilitate creation while maintaining the quality of the outputs on the platform.

Benchmarking AI Co-creation in Storytelling. The analysis of the state of the art in Human-AI co-creation must be complemented with a comparative evaluation of existing tools on the market. This benchmarking is crucial for identifying gaps and positioning the development of the Polariscope assistant. In this study, AI tools were evaluated, divided into two main focuses: fiction creation and contextual rigor. Tools such as Sudowrite, AI Dungeon, NovelAI, and Artflow, although excellent for producing fiction content, revealed a critical limitation in the Polariscope context: the absence of support for uploading and remixing large volumes of user-generated content. In contrast, tools focused on rigor and support for existing work were evaluated. For example, Grammarly, while integrating text generation features, maintains its primary focus on polishing and correcting writing. Canva Magic Write was examined for its ease in generating multimodal content for social networks, demonstrating the ability to create factual outputs and integrate various media (such as photos), depending on the given prompt. Based on this benchmarking, the market demonstrates a significant gap in solutions that combine the identified needs: a collaborative storytelling platform that integrates AI as an assistive tool for creators, focused on the use and enrichment of multimedia records from archive content. This result confirms the need to develop and evaluate a solution that fills this gap in the story co-creation process.

4 Methodology

This research is anchored in the Design-Based Research (DBR) methodology. DBR is chosen for its practical, problem-solving orientation, which aligns perfectly with the goal of developing and refining the AI assistant (the artifact) to improve the creative experience. The investigation is explanatory, seeking to answer the fundamental questions of why and how the introduction of the GenAI assistant impacts the quality and ease of the co-creation process. The DBR process is inherently interventionist (testing the artifact in a real-world setting) and iterative (using successive cycles of design, evaluation, and revision to refine the assistant). Hence, the following 4 phases were defined:

- **Phase 1.** State of the Art and Conceptualization (concepts; benchmarking of related work; personas; and functional requirements).
- **Phase 2.** Prototype Development (mockups and technical integration of services based on the conceptualized framework).
- **Phase 3.** Laboratory testing and Iteration (development and application of data collection instruments for evaluation of the prototype with users, followed by corrections and optimization).
- **Phase 4.** Field trial evaluation (development and application of a triangulation of data collection instruments, followed by final corrections and optimization).

The current paper presents the outcomes of Phases 1 and 2, detailing the conceptualized framework and its technical architecture.

For the upcoming evaluation phases, participants will be selected via convenience sampling and will include students from the University of Aveiro and students and teachers from a secondary school in Aveiro. We expect that these groups have experience with digital applications to avoid having accessibility barriers for standard App usage, so the research may focus on specific functionality. Two profiles are included: common users (to assess ease of creation) and experienced creators who are already familiar with the first version of Polariscope (no AI enhanced).

A mixed-methods approach will be used for data collection, conducted across Phase 3 and Phase 4. The data collection instruments are tailored to each evaluation phase: Phase 3 will focus on quick usability and functional feedback via short qualitative sessions within focus groups to assess usability issues, workflow flaws, and acute design needs, informing immediate interface adjustments. Phase 4 will involve a user experience assessment using remote in-app questions and quantitative questionnaires applied at scale, including validated instruments like the System Usability Scale (SUS) and AttrakDiff. This will be complemented by in-depth qualitative interviews with a group of participants for triangulation, which will explore the creators' motivations, sense of ownership, and the perceived quality of the stories to provide a comprehensive understanding of the results.

5 Proposed Solution

To move beyond a purely conceptual proposition, the proposed framework establishes a concrete technical architecture for integrating Generative AI into the Polariscope platform. The solution is designed to assist creators dynamically at the point of need. The overall architecture is structured into three interdependent layers: the Client Interface (Frontend), the Context Processing Engine (Backend), and the Large Language Model Inference (Cloud) (see Figure 1):

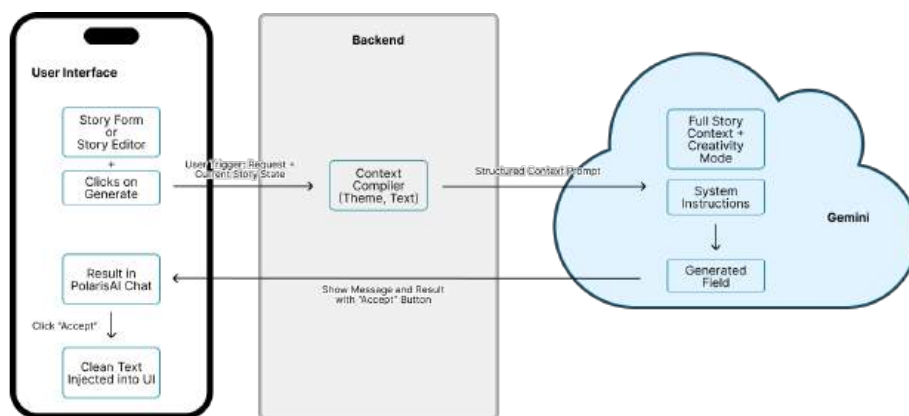


Fig. 1. Polariscope Architectural Workflow for AI-Assisted Story Co-Creation.

From a technical standpoint, the intelligent assistant (PolarisAI) is powered by the Google Gemini API, specifically utilizing the gemini-3-flash-preview model. To ensure the AI acts as a contextual co-author rather than a generic chatbot, a dedicated backend module handles strict prompt engineering. The framework integrates seamlessly into the user workflow by providing the following core capabilities as shown by Figure 2:

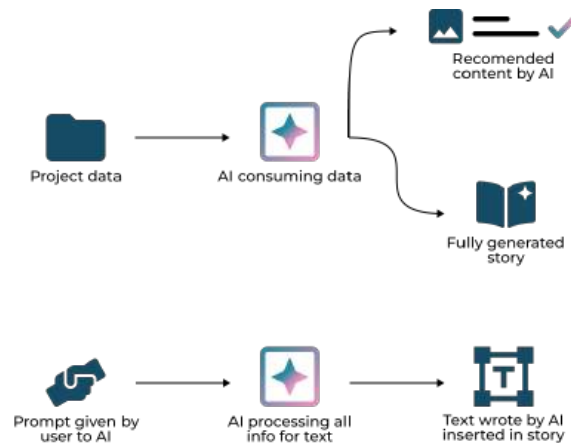


Fig. 2. Polariscope AI assistant Key Functionalities and their Workflow.

Intelligent Text Assistance. To prevent writer’s block while filling in text fields (titles, descriptions, or story paragraphs), the architecture defines two interaction paradigms. First, in-context quick buttons dynamically switch between "Generate" and "Rewrite" depending on whether a field is empty, allowing single-click assistance. Second, a dedicated chat interface allows users to insert specific instructions. When triggered, the backend dynamically compiles the user’s current progress (including the project’s name, available metadata of the contents, and previously written text blocks), into a structured context prompt. To guarantee that the AI’s output can be easily integrated into the platform’s UI, strict system instructions force the model to return responses using predefined structural tags (e.g., [[TITLE]], [[TEXT]]). A custom parser in the frontend processes these tags and presents the suggestions cleanly to the user as interactive cards with an “Accept” button that can be injectes the result directly into the story.

Content Recommendation. The AI scans the metadata of existing project content to identify assets that semantically align with the story the creator is currently writing. When the user selects the “Add Content” button, the backend matches the ongoing narrative context with the database archives, ensuring that the top multimedia contents deemed most relevant by the AI appear prioritized, bridging the gap between text generation and visual storytelling.

Full Story Generation. For projects with a substantial volume of content, able to be processed by GenAI to create a narrative, a dedicated button will be available on the

main project page. This function enables the project creator to automatically generate a complete story with minimal effort. However, this feature requires further assessment and evaluation.

This architectural framework ensures that the AI acts as a supportive layer. The prototype following these technical specifications is currently under development so that the research team can engage participants in the subsequent evaluation and optimization stages outlined in the methodology.

6 Conclusion

This paper presented a technical and conceptual framework for AI-assisted storytelling, aimed at addressing a core challenge in collaborative digital storytelling platforms: supporting users in transforming distributed multimedia contributions into coherent and meaningful narratives. Within the context of the Polariscope platform, the research explored how GenAI can be integrated as an assistive technology through a dedicated architecture to reduce barriers to story creation, such as limited time, creative confidence, and writer's block, while preserving human authorship and creative control.

The proposed solution positions AI as a creative partner within a Human-AI co-creation paradigm rather than as an autonomous content producer. The specified architecture defines a clear data flow involving context compilation, multimodal model inference via the Gemini API, and UI integration through custom parsing. Three complementary functionalities were established, targeting specific needs in the creative workflow: content recommendation, intelligent text assistance, and full story generation. By embedding AI assistance at moments of need through structured backend prompts, the framework seeks to increase efficiency, engagement, and the perceived quality of collaboratively created stories.

A central contribution of a prior assessment regarding automatic story creation [1][2] lies in the recognition that creativity in AI-generated narratives is highly context-dependent. Different story types require varying degrees of generative freedom, making adaptive control of AI parameters essential. This insight, combined with an updated benchmarking analysis of existing tools, revealed a clear gap in current solutions: the absence of platforms that effectively integrate AI-driven narrative support with the structured reuse of collective multimedia archives.

The research adopts a Design-Based Research methodology to ensure that technological development is closely aligned with iterative evaluation. With the architectural framework and initial prototype defined, future work will focus on the execution of Phase 3 and Phase 4. Controlled laboratory testing will be conducted to assess acute usability needs and workflow efficiency. Subsequently, a field trial will be conducted using real projects developed within the Polariscope platform by students of the University of Aveiro and the community of a secondary school in Aveiro. This trial will

evaluate the assistant's impact on story co-creation scenarios, providing empirical evidence of its effectiveness and receptivity to inform the final refinement of the platform.

Acknowledgments. This work is financially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project 2022.04424.PTDC and the UIDB/05460/2020.

References

1. Almeida, P., Teixeira, A., Velhinho, A., Raposo, R., Silva, T., Pedro, L.: Remixing and repurposing cultural heritage archives through a collaborative and AI-generated storytelling digital platform. In: *Proc. ACM Int. Conf. Interactive Media Experiences Workshops (IMXw 2024)*, pp. 100–104. ACM, New York (2024). <https://doi.org/10.1145/3672406.3672419>
2. Almeida, P., Raposo, R., Silva, T., Pedro, L., Velhinho, A., Alves, M., Roque, J., Graça, R.: AI-powered storytelling based on cultural heritage archives and collective memories. Poster presented at *A Inteligência Artificial no Futuro do Ensino Superior: Desafios e Oportunidades – 4º Encontro Aberto do Conselho Geral da Universidade de Aveiro*, Aveiro, Portugal (2025).
3. Nabirye, K.: Digital storytelling: Transforming narratives in the 21st century. *East Eur. J.* **6**(2), 35 (2025).
4. Trichopoulos, G., Alexandridis, G., Caridakis, G.: A survey on computational and emergent digital storytelling. *Heritage* **6**(2), 1227–1263 (2023). <https://doi.org/10.3390/heritage6020068>
5. Lomas, C., Burke, M., Page, C.L.: Collaboration tools (2008).
6. Lee, S., Bozeman, B.: The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* **35**(5), 673–702 (2005). <https://doi.org/10.1177/0306312705052359>
7. Karppinen, K., Koskela, K., Magnusson, C., Nore, V.: Experiences of online co-creation with end users of cloud services. In: *Lecture Notes in Computer Science*, vol. 6949, pp. 446–449. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23768-3_54
8. Herrmann, T.: Design issues for supporting collaborative creativity. In: *Proceedings of the European Conference on Computer-Supported Cooperative Work*, pp. 179–192 (2008).
9. Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P.S., Sun, L.: A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT. *J. ACM* **37**(111) (2023).
10. Du, H., Niyato, D., Kang, J., Xiong, Z., Zhang, P., Cui, S., Shen, X., Mao, S., Han, Z., Jamalipour, A., Poor, H.V., Kim, D.I.: The age of generative AI and AI-generated everything. *IEEE Netw.* **38**(6), 501–512 (2024). <https://doi.org/10.1109/MNET.2024.3422241>
11. Chakraborty, S., Björk, J., Dahlqvist, M., Rosen, J., Heintz, F.: A survey of AI-supported materials informatics. *Comput. Sci. Rev.* **59** (2026). <https://doi.org/10.1016/j.cosrev.2025.100845>
12. Shivaani, V.P., Muthupandi, G., Pavithra, S.: A review of generative AI in recommendation systems. In: *Proc. Int. Conf. Data Science, Agents and Artificial Intelligence (ICDSAAI 2025)* (2025). <https://doi.org/10.1109/ICDSAAI65575.2025.11011809>
13. Gozalo-Brizuela, R., Merchan, E.E.G.: A survey of generative AI applications. *J. Comput. Sci.* **20**(8), 801–818 (2023). <https://doi.org/10.3844/jcssp.2024.801.818>
14. Ng, D.T.K., Chan, E.K.C., Lo, C.K.: Opportunities, challenges and school strategies for integrating generative AI in education. *Comput. Educ.: Artif. Intell.* **8** (2025). <https://doi.org/10.1016/j.caeai.2025.100373>

15. Chan, C.K.Y., Colloton, T.: Generative AI in higher education. In: *Generative AI in Higher Education: The ChatGPT Effect*, p. 287. Routledge (2024). <https://doi.org/10.4324/9781003459026>
16. Hua, Y., Niu, S., Cai, J., Chilton, L.B., Heuer, H., Wohn, D.Y.: Generative AI in user-generated content. In: *Proc. CHI Conf. Human Factors in Computing Systems* (2024). <https://doi.org/10.1145/3613905.3636315>
17. Lyu, Y., Zhang, H., Niu, S., Cai, J.: A preliminary exploration of YouTubers' use of generative AI in content creation. In: *Proc. CHI Conf. Human Factors in Computing Systems*, p. 7 (2024). <https://doi.org/10.1145/3613905.3651057>
18. Haase, J., Pokutta, S.: Human-AI co-creativity: Exploring synergies across levels of creative collaboration. arXiv:2411.12527 (2024).
19. Heyman, J.L., Rick, S.R., Giacomelli, G., Wen, H., Laubacher, R., Taubenslag, N., Knicker, M., Jeddi, Y., Ragupathy, P., Curhan, J., Malone, T.: Supermind ideator: How scaffolding human-AI collaboration can increase creativity. In: *Proc. ACM Collective Intelligence Conf. (CI 2024)*, pp. 18–28 (2024). <https://doi.org/10.1145/3643562.3672611>
20. Malakar, P., Leeladharan, M.: Generative AI tools for collaborative content creation: A comparative analysis. *DESIDOC J. Libr. Inf. Technol.* **44**(3), 151–157 (2024). <https://doi.org/10.14429/djlit.44.03.19698>

Do Facto ao Acontecimento: proposta de arquitetura narrativa interativa para o jornalismo digital

Carlos Queirós¹[0000-0001-5911-617X] and Helder Caixinha²[0000-0002-7288-3290]

¹ DigiMedia, Department of Communication and Art, University of Aveiro, Portugal
carlosqueiros@ua.pt

² DigiMedia, Department of Communication and Art, University of Aveiro, Portugal
caixinha@ua.pt

Resumo. O baixo envolvimento dos jovens adultos no consumo de conteúdos noticiosos digitais tem sido associado à fragmentação informativa, à superficialidade do consumo e à inadequação de modelos narrativos lineares em ambientes interativos. Este artigo propõe uma arquitetura narrativa interativa de base teórica para o jornalismo digital, centrada na integração da agência do utilizador, sem comprometer a coerência narrativa, nem a intencionalidade editorial. A proposta assenta na distinção entre Facto e Acontecimento, conceptualizando o Acontecimento como uma construção discursiva dinâmica, mediada por interação. Articulando contributos da narrativa interativa, do *design* de informação e do jornalismo digital, o modelo define o jornalista como *Cyberbard*, responsável por orquestrar a arquitetura narrativa que regula a progressão do Acontecimento. A arquitetura proposta organiza-se segundo uma lógica de modularidade informativa em três camadas, permitindo diferentes níveis de envolvimento, aprofundamento e verificação. Introduce-se, ainda, o conceito de acesso fractal como mecânica de distribuição de elementos multimodais, adaptada a ecossistemas mediáticos. Esta abordagem contribui para o *design* de narrativas interativas, aplicadas a contextos noticiosos, promovendo experiências informativas mais coerentes, exploratórias e participativas.

Keywords: Narrativas interativas, Jornalismo digital, Jovens adultos, Modularidade informativa, Acesso fractal, Agência do utilizador

1 Introdução

Nas últimas décadas, a transformação digital reconfigurou a produção, circulação e consumo noticioso, alterando a relação dos jovens adultos com a informação jornalística [1]. Esta transformação é marcada pela convergência mediática descrita por Jenkins [2], em que conteúdos, práticas e significados circulam entre múltiplas plataformas, coexistindo paradoxalmente com a fragmentação dos ambientes mediáticos [3]. Esta dupla dinâmica redefiniu a interação entre utilizadores, conteúdos e instituições, fomentando expectativas de experiências interativas, multimodais e personalizadas [4], [5]. Neste contexto, os ambientes interativos podem ser lidos à luz da “inteligência coletiva” proposta por Lévy, segundo a qual o conhecimento circula e se recombina socialmente, beneficiando da partilha distribuída de saberes, tal como sublinha Hayek [6].

Para as gerações mais jovens, o consumo noticioso ocorre predominantemente em fluxos distribuídos (ou atomizados): redes sociais, sistemas algorítmicos e formatos

incidentais, afastando-se de modelos lineares e unidirecionais [1], [4], [7], [8]. Estes padrões de consumo reclamam novas formas de mediação jornalística adequadas às lógicas de atenção e exploração próprias deste segmento. Feio refere que a modalidade de mediação é decisiva para captar os jovens: o texto longo tende a ser evitado em favor de vídeos curtos, privilegiados pelas plataformas e pelas audiências [7]. Assim, as práticas mediáticas das audiências jovens reforçam a necessidade de conceber uma arquitetura narrativa interativa para conteúdos noticiosos, ajustada às lógicas participativas e às expectativas cognitivas deste público. Estas adaptações devem, contudo, ocorrer sem descurar a função central do jornalismo: a construção da realidade social a partir do Acontecimento [9]. A notícia não é a simples reprodução do Facto, mas o resultado de um processo discursivo de seleção e atribuição de significado [10], [9]. Na era digital a circulação fragmentada e intermitente da informação exige modelos narrativos capazes de representar e mediar a complexidade da realidade social.

Embora a literatura descreva amplamente a fragmentação do consumo noticioso e o papel da interatividade, subsiste uma lacuna ao nível do *design* narrativo: a ausência de arquiteturas narrativas interativas que articulem, num mesmo sistema, agência do utilizador, coerência narrativa e intencionalidade jornalística. Para responder a esta lacuna, o presente artigo propõe uma arquitetura narrativa interativa de base teórica aplicada ao jornalismo digital.

Este estudo insere-se num projeto de investigação no âmbito de um programa doutoral em curso sobre narrativas interativas aplicadas ao jornalismo digital. O público-alvo corresponde aos jovens adultos em Portugal, segmento populacional onde se observam mudanças significativas nos modos de consumo de conteúdos noticiosos.

2 A Problemática

Os dados do *Digital News Report 2025* evidenciam um afastamento dos jovens adultos dos meios noticiosos tradicionais. Nos 48 países analisados, (os indivíduos dos 18 aos 34 anos) o consumo digital afirma-se como o meio dominante, utilizado por cerca de 80% da população, contrastando com a utilização residual da televisão (25%), dos jornais impressos e da rádio (entre 3 e 5%). Em contrapartida, o acesso à informação ocorre maioritariamente através de fluxos mediados por algoritmos, e não por via direta dos *websites* noticiosos. Redes sociais e plataformas de vídeo, como Instagram, TikTok e YouTube, bem como o Facebook¹, apesar do seu declínio, constituem a principal fonte de notícias para uma parcela significativa deste público, a par de motores de busca e agregadores móveis, como o Google News ou o Apple News. Este padrão de acesso, marcado pelas plataformas digitais e consumo incidental, fornece a base empírica para a problemática da fragmentação e superficialidade informativa [5].

Este cenário configura o problema central da presente investigação, contudo, não se manifesta de forma isolada, pois articula-se com um conjunto de fatores estruturais que ampliam e justificam o afastamento dos jovens do consumo noticioso [11]. Para compreender este fenómeno é fundamental criar estratégias de aproximação dos jovens. A

¹ O relatório associa esta diminuição à reconfiguração estratégica da Meta ("Platform Reset"), que reduziu (para apenas 3%) a visibilidade de conteúdos noticiosos nos *feeds*, a favor de formatos de entretenimento e vídeo curto como resposta ao crescimento do TikTok [5].

problemática desta investigação organiza-se em cinco dimensões teóricas, quatro delas resultam predominantemente de lógicas algorítmicas: a atomização do conteúdo, a superficialidade do consumo, a descontextualização e a erosão da coerência narrativa, enquanto a quinta diz respeito à persistência de um discurso monológico associado à narrativa linear tradicional.

1. Atomização do conteúdo noticioso: no ecossistema digital observa-se um agravamento do *news gap* descrito por Boczkowski e Mitchelstein, uma vez que os padrões de consumo se afastam das hierarquias editoriais e convergem para micro-unidades fragmentadas de informação [12]. Neste contexto, a notícia deixa de operar como produto acabado e passa a circular, nos termos de Bruns, como um fluxo atomizado, no qual o utilizador assume funções de curadoria ao recombina-los fragmentos recontextualizados por algoritmos e plataformas sociais [13];
2. A superficialidade do consumo: a lógica fragmentada do ambiente digital privilegia a brevidade em detrimento da profundidade. Patterson, citado por Dimitrova, alerta que o jornalismo *online* favorece formatos curtos e emocionalmente carregados, orientados para captar atenção momentânea e não para informar de forma sustentada [14]. Este efeito é reforçado pelo acesso incidental: Fletcher e Nielsen demonstram que a exposição a notícias nas redes sociais gera *awareness*, mas raramente conduz ao aprofundamento, resultando num conhecimento superficial da realidade [15];
3. Descontextualização: a lógica do *feed* contínuo introduz um problema de sintaxe informativa que compromete a articulação narrativa. Hermida descreve o *ambient journalism*² como um regime de atualizações sucessivas que produz observações fragmentadas em detrimento de narrativas estruturadas [4]. Esta condição é aprofundada pela teoria da *deep mediatization*, de Couldry e Hepp (citados por Lewis), segundo a qual as infraestruturas digitais reorganizam os Acontecimentos em unidades dispersas, enfraquecendo a capacidade de estabelecer relações temporais e causais [15]. No consumo via redes sociais, este défice de contexto torna-se particularmente evidente, como demonstra Bode ao identificar a ausência recorrente de pistas contextuais essenciais à interpretação da informação [16];
4. Erosão da coerência narrativa: as carências anteriormente identificadas acumulam-se e produzem um efeito estrutural sobre a forma narrativa jornalística. Jacobson, Marino e Gutsche argumentam que a lógica não linear do hipertexto e da fragmentação multimédia perturba a coerência narrativa, substituindo o *storytelling* sustentado por estruturas modulares e descontínuas [17]. Em ecossistemas dominados por atualizações breves e sistemas de recomendação que desorganizam a cronologia dos Factos, a progressão narrativa torna-se difícil de manter, como observam Dowling e Vogan [18];
5. Discurso monológico: para além dos condicionamentos algorítmicos, persiste um desajuste na matriz discursiva do jornalismo tradicional, que continua a operar numa perspetiva monológica e unidirecional. À luz de Bakhtin, este modelo assenta na imposição de uma voz de autoridade que não convoca a resposta do interlocutor,

² Hermida utiliza o termo “*ambient journalism*” como um sistema de *awareness* baseado na circulação contínua de fragmentos informativos, característica das plataformas digitais [4].

relegando o público para uma posição passiva [19]. Como enfatiza Deuze [20], o jornalismo digital dos *mass media* tradicionais manteve uma cultura de publicação (*Broadcasting* – um para muitos) em detrimento da participação, falhando a promessa de se tornar verdadeiramente dialógico [21]. Esta lógica colide com as expectativas das audiências jovens, habituadas à Cultura da Convergência descrita por Jenkins [2].

Estas cinco dimensões evidenciam que os desafios do consumo noticioso não são apenas de natureza editorial ou tecnológica. Decorrem também da ausência de arquiteturas narrativas capazes de organizar estas dinâmicas numa progressão coerente. O problema assume, portanto, uma natureza estrutural: trata-se de conceber formas narrativas capazes de articular múltiplas unidades informativas sem perder continuidade interpretativa. Neste sentido, a modularidade descrita por Manovich [22] constitui um argumento fundamental: se a informação digital é estruturalmente modular, então uma arquitetura narrativa jornalística adequada ao meio não se deve limitar a adaptar a forma tradicional da notícia, mas exige uma organização narrativa modular, baseada em unidades informativas recombinaíveis e escaláveis e capazes de acomodar a fragmentação do consumo e a necessidade de coerência narrativa [23], [24], [25].

3 Do Facto ao Acontecimento: Proposta de uma Arquitetura Narrativa Interativa

A resposta aos desafios identificados na secção anterior reside na conceção de uma arquitetura narrativa capaz de estruturar a informação digital de forma coerente e explorável, permitindo articular múltiplos módulos informativos sem comprometer a continuidade interpretativa do Acontecimento. Neste contexto, propõe-se uma Arquitetura Narrativa Interativa que articula a ontologia jornalística com o *design* narrativo, convertendo o leitor em *interactor*³, capaz de navegar dentro de um sistema com vários módulos informativos [5], [9].

Para desenhar uma nova arquitetura narrativa interativa é fundamental compreender, primeiro, o processo de transformação da “matéria-prima” em “produto final”. Afastando-se de uma visão mecanicista, este modelo assume que o jornalismo não reflete a realidade, mas constrói o que Muniz Sodré (citado por Carvalho) [21] define como “efeitos de real” [21]. Não se trata de ficcionar o real, mas de o configurar. Seguindo a perspectiva de Seixas [9], o jornalismo não opera diretamente sobre a realidade bruta, mas sim através de um processo de seleção e codificação que pode ser decomposto em três níveis distintos:

1. A Realidade e o Facto - a realidade apresenta-se como um fluxo contínuo e caótico de ocorrências; o Facto é o recorte inicial dessa realidade, é a ocorrência verificável, mas que, no seu estado bruto, permanece muda, dispersa e desprovida de sentido

³ **Interactor** – é um termo utilizado na teoria das narrativas interativas para designar o participante que interage com o sistema narrativo, distinguindo-se do leitor ou espetador das narrativas lineares. Ao contrário destes, o *interactor* participa ativamente na progressão da narrativa através de escolhas ou ações que influenciam o percurso dentro do espaço narrativo. Esta distinção é recorrente na literatura sobre narrativas digitais e interativas [23], [25].

social até ser observada [9]. Numa fase inicial, o Facto apresenta-se como um “candidato” para fazer parte de um Acontecimento.

2. O Objeto de Realidade - é nesta fase que intervém a competência do jornalista; ao seleccionar um Facto com base em critérios de noticiabilidade (relevância, interesse público, impacto), transforma-o num “Objeto de Realidade”. Este objeto é a matéria-prima processada, retirada do fluxo contínuo e preparada para ser trabalhada. Ao fazer isto, estabelece o que Alsina chama de um “Mundo de Referência”: um recorte validado e verosímil que serve de base para a história [26]. Seixas descreve esta fase como a transformação da matéria-prima bruta em algo processável [9].
3. O Acontecimento - o estágio final é o Acontecimento; diferente do Facto (que pertence ao real e ao passado), o Acontecimento pertence ao discurso. É, na visão de Seixas, uma configuração simbólica onde se atribui sentido, contexto e memória; é uma construção narrativa completa com uma configuração dotada de memória e sentido social [9]. Alsina expande esta visão chamando-lhe um “Mundo Possível”: uma construção narrativa que, através de marcas de credibilidade, cria a ilusão referencial necessária para que o público aceite a história como verdadeira [26].

Em síntese, esta conceção de mediação parte do princípio de que o jornalismo não se limita a refletir a realidade, mas participa na sua construção discursiva. Como observa Muniz Sodré (citado em [21]), o discurso jornalístico produz “efeitos do real”, isto é, configurações simbólicas que procuram apresentar-se como referências da realidade social.

Neste enquadramento, Seixas propõe compreender este processo através de três níveis analíticos [9]. A realidade apresenta-se inicialmente como um fluxo contínuo do qual o Facto surge como um recorte verificável, ainda desprovido de significado narrativo. Através dos processos de seleção e noticiabilidade, esse recorte transforma-se em Objeto da Realidade, tornando-se uma unidade processável no sistema editorial jornalístico e contribuindo para a constituição de um “Mundo de Referência” socialmente validado [20]. O processo culmina no Acontecimento, que já não pertence ao plano do real enquanto tal, mas ao domínio do discurso: uma configuração simbólica que organiza os factos num “Mundo Possível” sustentado por marcas de credibilidade junto do público [9], [20].

Esta distinção entre níveis de mediação evidencia que o Acontecimento noticioso resulta de um processo de construção discursiva, princípio que será retomado na secção seguinte para enquadrar a organização modular da informação no contexto das narrativas digitais interativas.

3.1 O Jornalista como *Cyberbard*: a *Proto-story*⁴

A transição para a interatividade, entendida como a passagem de uma narrativa linear para estruturas que incorporam a agência do utilizador, transforma o papel do jornalista;

⁴ **Proto-story** - é um conceito desenvolvido por Hartmut Koenitz no âmbito das *Interactive Digital Narratives (IDN)*, refere-se ao conjunto total de elementos narrativos e possibilidades de percurso concebidos pelo autor no sistema narrativo. Trata-se, portanto, da estrutura subjacente que define o espaço de histórias potencialmente experienciáveis pelo utilizador durante a interação.

este deixa de ser apenas autor de um texto fixo e assume funções de *designer de narrativas* [20] ou *Cyberbard*⁵ [27], responsável por projetar e orquestrar o sistema de interação entre utilizador e conteúdo. Neste contexto, a estrutura narrativa torna-se determinante. Como observa Nissen, na Era da Informação, a organização estrutural da informação assume um papel determinante [28], exigindo a necessidade de unificar os diferentes elementos informativos (multimodais). Esta exigência conduz à arquitetura de uma estrutura narrativa integradora, próxima do conceito de *proto-story* proposto por Koenitz, capaz de articular múltiplos módulos informativos e evitar a dispersão interpretativa [29].

O jornalista - arquiteto que opera sob quatro princípios: (a) Modularidade, definindo unidades informativas multimodais e sua interligação; (b) Granularidade, estabelecendo níveis de profundidade para leituras em camadas; (c) Interconectividade, a arquitetura semântica que liga as partes ao todo, promovendo uma “narrativa emergente” onde o sentido resulta da interação dinâmica; e (d) Navegação, que delimita caminhos para resolver o Paradoxo Narrativo (tensão entre agência do utilizador e coerência autoral). Ao desenhar estes limites, o *Cyberbard* cria um “espaço de possibilidades” onde a liberdade de navegação é exercida sem comprometer a integridade editorial do Acontecimento [30].

3.2 Proposta de Arquitetura: a Modularidade Informativa

Partindo da “Pirâmide Deitada”⁶ de Canavilhas [31], [32], a Modularidade Informativa adapta a navegação por níveis às condições estruturais do ambiente digital contemporâneo. Prefere-se o termo “Módulo” em detrimento do “bloco”, para alinhar com a modularidade proposta por Manovich [22]: uma “estrutura fractal” onde os elementos mantêm dependência funcional. Assumindo-se como unidade base que distingue ontologicamente Facto de Acontecimento, o módulo organiza-se internamente em camadas para conciliar *awareness* e compreensão. Esta estrutura, gerida pelo *Cyberbard*, permite ao utilizador exercer agência para penetrar a superfície noticiosa através de três camadas (C1, C2 e C3), reconstruindo a complexidade da realidade social:

1. **C1 - O Marco Noticioso (Lead – o conhecimento):** esta camada atualiza a função do *Lead*, servindo como porta de entrada nos *feeds* de redes sociais e fluxos de consumo imediato. Baseada na distinção cognitiva de Majin [33], a C1 satisfaz a necessidade

⁵ *Cyberbard* - conceito cunhado por Janet Murray [27] para designar o autor de sistemas narrativos operacionalizado na era digital. No contexto da Arquitetura Narrativa Interativa proposta neste artigo, este conceito é operacionalizado no domínio do jornalismo digital para descrever o papel do editor, jornalista ou *designer* responsável pela configuração do espaço narrativo. Neste sentido, o *Cyberbard* não escreve apenas um texto fixo, mas projeta um sistema interativo: define os nós narrativos, seleciona as vozes de relevância social, determina os dados contextuais pertinentes e estabelece os limites do espaço de exploração disponível ao *interactor*.

⁶ A “Pirâmide Deitada” proposta por Canavilhas organiza a notícia digital em níveis de aprofundamento informativo acessíveis através de navegação hipertextual. A utilização do termo “módulo” neste trabalho aproxima-se, contudo, da noção de modularidade dos novos média descrita por Manovich, entendida como a organização da informação em unidades recombinaíveis.

de "Conhecimento" (*Knowledge*), isto é, a recolha factual que precede a explicação. O foco é responder às perguntas fundamentais (Quem? O Quê? Quando? Onde? Como?), funcionando como um "alerta" do espaço temporal. Ao isolar o Facto, o sistema respeita a "flecha do jornalismo"⁷, evitando que a opinião contamine a primeira impressão da realidade [33].

2. **C2 – O Aprofundamento** (vozes e perspectivas – o caminho para a compreensão): ativada pela agência do utilizador, esta camada opera a transição do conhecimento (do Facto) para a "Compreensão" (*Understanding*). Para Majin, esta camada implica detetar padrões de causa-efeito e responder ao "Porquê?" [33]. Revelando a complexidade do *Objeto de Realidade*, a C2 abandona a voz monológica por uma estrutura polifónica (entrevistas, contraditório, nuances interpretativas). Ao dissociar a explicação dos Factos (C1), o modelo mitiga o "jornalismo liderado pela narrativa", permitindo que o sentido seja uma construção ativa do leitor através da pluralidade de perspectivas, em vez de uma receção passiva.
3. **C3 - A Base Documental** (contexto e dados – a verificação): esta é a camada de imersão e memória, constitui a âncora de evidência empírica do sistema, agregando dados brutos, relatórios e documentos originais. A sua função primordial é combater o "*factinon*"⁸ de Majin para o híbrido epistémico que funde, indistintamente, Factos e opiniões, resistindo à verificação [33]. Aqui, o utilizador valida se a narrativa (C2) corresponde aos dados reais, exercendo um controlo democrático sobre a informação e acedendo às causas estruturais mais profundas que transcendem a atualidade imediata.

A Modularidade Informativa é uma construção dinâmica guiada pelo *Cyberbard*, esta arquitetura acompanha o ciclo noticioso: inicia-se na C1, que responde à urgência do *breaking news*, acoplando as camadas subsequentes (C2 e C3) como atualizações incrementais. O Acontecimento é construído em tempo real, do Facto isolado à complexidade contextual, respeitando a velocidade (editorial) digital sem comprometer o rigor da verificação final, dispensando a exigência inicial da "história completa". A arquitetura é elástica: as camadas C2 e C3 são contingentes, sendo ativadas apenas por necessidade editorial de aprofundamento, adaptando, assim, a densidade informativa à complexidade de cada acto.

3.3 Mecânicas de Distribuição: o Acesso Fractal

Para maximizar o alcance do jornalismo na economia da atenção, esta arquitetura prevê que a modularidade informativa ultrapasse o modelo tradicional de *site* de notícias. Propõe-se uma lógica de acesso fractal, inspirada na estrutura modular descrita por Manovich (2002), em que elementos multimodais de elevado valor semântico, como

⁷ A metáfora da "Flecha do Jornalismo", de Majin, prevê que o processo noticioso obedeça a uma direção: partir da evidência factual para a conclusão. Se começar pela conclusão e seleccionar apenas os Factos que a servem resulta no "Jornalismo Liderado pela Narrativa", comprometendo a isenção e a verdade [33].

⁸ Majin designa "*Factinon*" como algo "híbrido epistémico" que funde indistintamente Factos e opiniões. Ao operar no domínio da crença, este fenómeno resiste ao *fact-checking* ao tratar opiniões como Factos irrefutáveis e Factos como pontos de vista [33].

vídeos de atualização (C2) ou infografias (C3), possam ser atomizados e distribuídos de forma autónoma em *feeds* de redes sociais, funcionando como âncoras de envolvimento [22]. Para evitar os riscos de descontextualização inerentes a este tipo de distribuição, o sistema narrativo integra uma lógica de encaminhamento condicional, orientada pelo histórico de navegação do utilizador. Definem-se dois cenários:

- (a) contextualização - para utilizadores que interagem pela primeira vez, o clique redireciona para C1, garantindo o contacto inicial com o Lead (Conhecimento) antes do acesso ao elemento multimodal;
- (b) acesso direto - para utilizadores recorrentes que já consultaram a C1, o sistema permite um atalho para as atualizações específicas nas camadas C2 ou C3.

4 Conclusão e Defesa da Proposta

A proposta parte do diagnóstico de um afastamento significativo dos jovens adultos face aos formatos noticiosos dos *mass media* tradicionais. Identificou-se que este afastamento não resulta apenas da migração tecnológica, mas de uma mudança estrutural nas formas de atenção, interação e atribuição de sentido. Neste contexto, a proposta - Arquitetura Narrativa Interativa, defende que a resposta jornalística deve passar por uma reconfiguração discursiva que incorpore a agência do utilizador, a modularidade informativa e uma mediação narrativa mais adaptativa [5].

Esta arquitetura organiza-se em torno da articulação entre Facto e Acontecimento, reconhecendo que a realidade jornalística é uma construção discursiva que exige novos modelos representacionais no ecossistema digital. A introdução do conceito de *proto-story*, operada pelo jornalista enquanto *Cyberbard*, permite transformar a fragmentação dos fluxos informativos num sistema coerente, navegável e verificável.

Com base em três camadas (C1, C2, C3), propõe-se uma arquitetura que assegura uma progressão respeitando, simultaneamente, a velocidade do consumo contemporâneo e a integridade editorial do jornalismo. O acesso Fractal, inspirado em Manovich, reforça a adaptabilidade do modelo às lógicas da atenção distribuída [22].

Esta abordagem contribui para uma teoria do jornalismo digital orientada para a participação crítica e a valorização do Acontecimento como estrutura com sentido. Ao integrar modularidade, interatividade e verificação, esta arquitetura narrativa visa reaproximar os jovens das práticas jornalísticas, abrindo caminho para novas formas de literacia mediática e cidadania digital.

O presente estudo tem, contudo, algumas limitações, na medida em que se circunscreve no plano teórico e de projeto: a arquitetura narrativa proposta, incluindo as mecânicas de distribuição, carecem, ainda, de implementação e validação empírica junto de utilizadores reais. Esta é uma etapa que constitui a continuidade natural desta investigação. Paralelamente, o desenvolvimento de uma arquitetura desta natureza deverá ter impacto na formação de editores-jornalistas, com funções em redações habituadas a uma produção de conteúdos lineares e que agora poderão ser confrontadas com uma nova realidade ao nível da tecnologia e da ontologia jornalística.

Agradecimentos

Esta investigação é desenvolvida no DigiMedia - Digital Media and Interaction Research Centre do Departamento de Comunicação e Arte - DeCA da Universidade de

Aveiro, no âmbito do Programa Doutoral em Novos Média. Este estudo conta com o apoio e a parceria da empresa ClusterMedia Labs. O 1º autor agradece à Fundação para a Ciência e a Tecnologia - FCT o financiamento da bolsa (bolsa 2025.04770.BDANA) para a realização desta investigação.

Referências

- [1] G. Cardoso, M. Paisana, and A. Pinto-Martinho, “Digital News Report Portugal 2025 Obercom - Observatório da Comunicação,” 2025. doi: [org/10.5281/zenodo.15599071](https://doi.org/10.5281/zenodo.15599071).
- [2] Henry Jenkins, *Cultura da Convergência*, 2 ed. São Paulo: NYU Press, 2009.
- [3] T. B. Ksiazek, “Fragmentation of News Audience,” *The International Encyclopedia of Journalism Studies*, pp. 1–4, Apr. 2019, doi: [10.1002/9781118841570.IEJS0049](https://doi.org/10.1002/9781118841570.IEJS0049).
- [4] A. Hermida, “Twittering The News,” *Journalism Practice*, vol. 4, no. 3, pp. 297–308, 2010, doi: [10.1080/17512781003640703](https://doi.org/10.1080/17512781003640703).
- [5] N. Newman, A. Ross Arguedas, C. T. Robertson, R. Kleis Nielsen, and R. Fletcher, “Reuters Institute Digital News Report 2025,” 2025. doi: [10.60625/risj-8qqf-jt36](https://doi.org/10.60625/risj-8qqf-jt36).
- [6] Hayek and F.A., *Law, legislation and liberty: A new statement of the liberal principles of justice and political*, Vols. 1-3. Routledge, 1993.
- [7] C. Feio and L. Oliveira, “Scrolling Through the Feed - How do young people in Portugal consume news on social media?,” *Observatorio*, pp. 1–15, 2025, doi: [10.15847/OBSOBS19520252670](https://doi.org/10.15847/OBSOBS19520252670).
- [8] M. Maneta, L. Torre, A. Rodrigues, I. Amaral, and P. Jerónimo, “Jovens e consumo mediático: uma análise comparativa entre desertos e não-desertos de notícias em Portugal,” *Media & Jornalismo*, vol. 24, no. 45, p. e4512, Nov. 2024, doi: [10.14195/2183-5462_45_12](https://doi.org/10.14195/2183-5462_45_12).
- [9] L. Seixas, *Redefinindo os Gêneros Jornalísticos: proposta de novos critérios de classificação*. Livros LabCom, 2009.
- [10] A. Roosvall, “Gaye Tuchman (1978) Making News: A Study in the Construction of Reality,” *Classics in Media Theory*, pp. 227–241, Jan. 2024, doi: [10.4324/9781003432272-17](https://doi.org/10.4324/9781003432272-17).
- [11] F. Gil, *Ciência e Filosofia*, USP., vol. N° 3. São Paulo: Universidade de São Paulo, 1986.
- [12] P. Boczkowski and E. Mitchelstein, “The News Gap: When the Information Preferences of the Media and the Public Diverge,” *Perspectives on Politics*, vol. 13, no. 1, pp. 221–223, Mar. 2015, doi: [10.1017/S1537592714003727](https://doi.org/10.1017/S1537592714003727).
- [13] J. Y. M. Nip, “Gatewatching and news curation: Journalism, social media, and the public sphere,” *Digital Journalism*, vol. 7, no. 4, pp. 543–544, Apr. 2019, doi: [10.1080/21670811.2019.1598883](https://doi.org/10.1080/21670811.2019.1598883).
- [14] D. V. Dimitrova, “Informing the News: The Need for Knowledge-Based Journalism by Thomas E. Patterson,” *Mass Commun. Soc.*, vol. 20, no. 1, pp. 136–139, Jan. 2017, doi: [10.1080/15205436.2016.1234163](https://doi.org/10.1080/15205436.2016.1234163).
- [15] W. W. Lewis, “The mediated construction of reality,” *Int. J. Perform. Arts Digit. Media*, vol. 16, no. 2, pp. 213–215, May 2020, doi: [10.1080/14794713.2020.1730071](https://doi.org/10.1080/14794713.2020.1730071).

- [16] L. Bode, "Political News in the News Feed: Learning Politics from Social Media," *Mass Commun. Soc.*, vol. 19, no. 1, pp. 24–48, Jan. 2016, doi: 10.1080/15205436.2015.1045149.
- [17] S. Jacobson, J. Marino, and R. E. Gutsche, "The digital animation of literary journalism," *Journalism*, vol. 17, no. 4, pp. 527–546, May 2016, doi: 10.1177/1464884914568079.
- [18] D. Dowling and T. Vogan, "Can We 'Snowfall' This?: Digital longform and the race for the tablet market," *Digital Journalism*, vol. 3, no. 2, pp. 209–224, Mar. 2015, doi: 10.1080/21670811.2014.930250.
- [19] M. Bakhtin, *Estética da Criação Verbal*, 2. ed. São Paulo: Martins Fontes Editora Ltd., 1997.
- [20] M. Deuze, "The web and its journalisms: Considering the consequences of different types of newsmedia online," *New Media Soc.*, vol. 5, no. 2, pp. 203–230, Jun. 2003, doi: 10.1177/1461444803005002004.
- [21] C. Carvalho, *A construção da notícia: interseções entre jornalismo e comunicação estratégica*. 2019. doi: org/10.7476/9786556301808.
- [22] L. Manovich, *The Language of New Media*. The MIT Press, 2002.
- [23] M. Ryan, K. Foote, and M. Azaryahu, *THEORY AND INTERPRETATION OF NARRATIVE*. 2016.
- [24] M.-L. Ryan, "My Narratology An Interview with," 2014.
- [25] H. Koenitz, J. Barbara, and M. P. Eladhari, "Interactive digital narrative (IDN)—new ways to represent complexity and facilitate digitally empowered citizens," *New Review of Hypermedia and Multimedia*, vol. 28, no. 3–4, pp. 76–96, Oct. 2023, doi: 10.1080/13614568.2023.2181503.
- [26] M. Alsina, "La produccion de la noticia," 1989, pp. 58–63.
- [27] J. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. The MIT Press, 1997.
- [28] A. M. Nissen, "The metanarrative paradigm," 2017. Accessed: Dec. 09, 2025. [Online]. Available: <https://researchportal.murdoch.edu.au/esploro/outputs/doctoral/The-metanarrative-paradigm/991005540172207891>
- [29] H. Koenitz, J. Barbara, and M. P. Eladhari, "Interactive digital narrative (IDN)—new ways to represent complexity and facilitate digitally empowered citizens," *New Review of Hypermedia and Multimedia*, vol. 28, no. 3–4, pp. 76–96, 2022, doi: 10.1080/13614568.2023.2181503.
- [30] J. Smed, T. " Bgt, " Suovuo, N. Trygg, and P. Skult, "Lecture Notes on Interactive Storytelling," 2022, Accessed: Dec. 10, 2025. [Online]. Available: <https://www.utu-pub.fi/handle/10024/159319>
- [31] J. Canavilhas, "Hipertextualidade Novas arquiteturas noticiosas," 2014.
- [32] J. Canavilhas, "Jornalismo na Web: Da Pirâmide Invertida à Pirâmide Deitada," *Aprender*, pp. 58–65, Jun. 2012, doi: 10.58041/APRENDER.113.
- [33] G. Majin, "What is a News Narrative?: Reversing Journalism's Arrow and the Problem of 'Factinion,'" *The Journal of Communication and Media Studies*, vol. 6, no. 2, pp. 57–70, 2021, doi: 10.18848/2470-9247/CGP/V06I02/57-70.

Personalized Multimedia Story Viewing Experience on the Polariscope Platform

Jónatas Roque^[0009-0004-5301-0947], Pedro Almeida^[0000-0001-5878-3317],
Telmo Silva^[0000-0001-9383-7659] and Ana Velhinho^[0000-0001-9978-8317]

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

Abstract. We are living through an era marked by the acceleration of digital content consumption, a hyper-stimulating ecosystem that demands adaptation strategies, and the personalization of digital storytelling experiences. This research addresses this issue in the context of digital narratives consumption. The paper, in the scope of the Polariscope project, details the development and evaluation of a user-centered approach for the visualization of digital narratives, which incorporates personalized viewing modes inspired by filtering and recommendation systems, with the goal of offering a personalized and customizable experience do Polariscope users. This project aims to qualitatively enhance the consumption experience, increasing user satisfaction. The methodology followed a user-centered and iterative structure combining qualitative and quantitative techniques, employing a Design-Based Research (DBR) framework with a quasi-experimental approach. This methodology resulted in five stages, initially the state-of-the-art and literature review, then the development of mockups and focus group with the target audience of the Polariscope platform, then a development phase, followed by laboratory tests to the same audience as in the focus group, and finally, the analysis of the data. This iterative process enabled the assessment of the solution's effectiveness by incorporating users' feedback. Based on the carried research, it was found that the personalized viewing mode functionality significantly improves the user experience and satisfaction in consuming digital narratives, while also contributing to the expected user loyalty. This user-centered and personalization-driven solution shows potential for application in other contexts, fostering innovation in digital storytelling.

Keywords: Adaptive Interfaces, Digital Narratives, Digital Storytelling, Personalization, User Experience Customization.

1 Introduction

The fast evolution of the Web into a participatory and socially driven ecosystem has altered user expectations regarding interactivity, personalization, and narrative consumption. Digital Storytelling (DS) has emerged as a powerful form of communication capable of strengthening emotional engagement and promoting immersive experiences. Digital storytelling has evolved significantly over the past decades and now requires

more diverse features and modes of narrative consumption to meet users' expectations. As audiences increasingly engage with complex, multimedia-rich environments, storytelling platforms must offer more flexible, immersive, and personalized ways of experiencing narratives in order to enhance engagement, comprehension, and emotional connection. [2] [3] [4]. This research investigates how personalized viewing modes can enhance the visualization and consumption of digital narratives within the Polariscope service [1], a platform designed for the co-creation and sharing of collective and individual memories. Polariscope operates as a Progressive Web App (PWA), centered around three core elements: Projects, Stories, and Contents. Stories are user-generated narrative structures composed of text, multimedia elements (contents), and user contributions in it. Projects aggregate Stories under a thematic umbrella related to cultural events or topics, while Contents (images, videos, audios and texts shared by users) provide contextual richness. The story page structure, within Polariscope, relied primarily on a traditional vertical scroll model for story visualization. To provide a more immersive and customized alternative to meet differentiated users' preferences, the research was designed to conceptualize and evaluate a new feature to enhance the experience of Digital Storytelling of the Stories of Polariscope.

Hence, this project was guided by the following research question: "What personalized viewing modes for Digital Storytelling (DS) are most valued for improving the user experience?" To answer this question, the study aims to:

- Identify relevant personalized viewing modes applicable to Digital Storytelling (DS) platforms.
- Analyse the user needs and expectations regarding narrative consumption modes.
- Develop and prototype a functional Story visualization system called *Play Mode*
- Validate the implemented system through iterative testing with real users.
- Assess the impact of personalization options on story immersion, comprehension, and viewing satisfaction.

The document is structured into six sections: the first presented the aim of the study; the second contextualizes the Polariscope platform and the foundations of the Play Mode; the third details the methodology and its phases; the fourth systematizes the obtained results; and the last draws the main findings and the pathway for future work.

2 The Polariscope platform

The Polariscope platform offers an environment where personal and collective memories are co-created and archived. Its Stories rely on textual and multimedia elements to convey emotional depth and contextual detail. Given the platform's collaborative and narrative-driven nature, the need for improved visualization techniques aligns strongly with theoretical discussions around participatory culture [5], collective intelligence [6], and the cognitive impact of Digital Storytelling [7]. Emotions and cognitive engagement are central to DS impact [8] [9].

Based on this data, the need to improve the way these digital narratives are consumed was identified, namely through, enhancing the interface navigation and layout, and

implementing personalization mechanisms to allow adaptability to user preferences/personalization. This led to the conceptualization of the *Play Mode*, designed as a novel feature to enhance how Stories are consumed in the Polariscope platform. The *Play Mode* presents each section of a Story in a segmented, sequential format like a carousel of slides, instead of the already available feed format. Each slide corresponds to a narrative element from the original Story, such as text, image, video, audio clip, text citation, or other type of sections, promoting a more focused and immersive consumption experience (see Figure 1). The interface is minimalist and dark-toned to reduce distractions and includes controls for pause/play, and navigation between slides. This approach contrasts with the traditional vertical scroll, offering a more cinematic and controlled experience. This structure aligns with narrative visualization frameworks analyzed during the benchmarking and literature review [10] [11] [12].

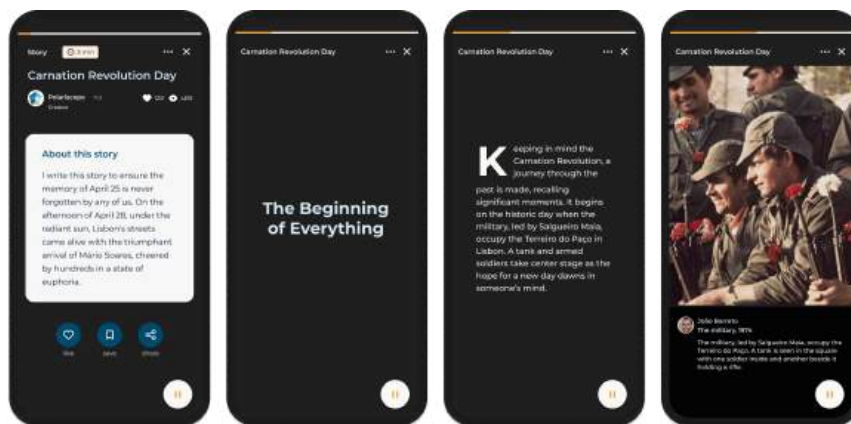


Fig. 1. Story Play Mode interface of the Polariscope platform.

3 Methodological approach

The methodology adopted in this study followed a user-centered and iterative research structure, combining qualitative and quantitative techniques to ensure a comprehensive understanding of user needs and experiences. The study was grounded in a Design-Based Research (DBR) framework, complemented by a quasi-experimental approach, allowing for the iterative development, testing, and refinement of the proposed solution in real usage contexts while systematically evaluating its impact on user experience. The process was divided into five key stages, ensuring continuous feedback integration and solution refinement. This methodological design follows mixed-methods principles [13] [14] and a Design-Based Research approach [15] [16].

3.1 Phase 1: Literature review and state-of-the-art

A comprehensive review of academic literature and market platforms established the theoretical and practical foundation for designing personalized viewing modes [1]. This

included an analysis of DS frameworks, personalization models, filtering systems (e.g., based on emotion, content type, context), and recommendation algorithms (collaborative filtering, content-based). Eight digital platforms were analyzed to identify common practices and innovative features, all in terms of viewing modes and also other similar features like personalization models, filters and recommendation algorithms: YouTube (video sharing), Spotify (music/podcast streaming), Blinkist (book summaries), SmartNews (news aggregation), Pinterest (visual discovery), Facebook (social networking), Museu da Pessoa - Memô (collective narrative museum), and Instagram (social media with Stories format). This analysis provided valuable insights for the *Play Mode* of the Polariscope platform, especially for the potential viewing modes to develop.

3.2 Phase 2: Prototype development and focus group evaluation.

Initial concepts for personalized viewing modes, derived from the previous phase, materialized into a medium-fidelity interactive prototype built in Figma (see Figure 1). This prototype simulated the *Play Mode*'s core interaction and showcased proposals of potential viewing modes (e.g., summary mode; by content type). This prototype served as the central discussion tool in a focus group [17] conducted with 6 participants. The sample was non-probabilistic and convenience-based, consisting of university students from various fields who were selected for their familiarity with digital technologies and narrative consumption habits, representing a key segment of Polariscope's target audience. The focus group was conducted as an intermediate and exploratory stage of the research process, with the goal of informing the construction and implementation of the *Play Mode* within the Polariscope platform. Initially, the session aimed to validate the proposed concepts related to the *Play Mode* functionality and to collect participants' preferences regarding which viewing modes were considered the most relevant and useful. The qualitative insights gathered during this phase directly supported the decision-making process and guided the subsequent development of the *Play Mode* prototype integrated into the platform.

3.3 Phase 3: Platform implementation.

Guided by the feedback from the focus group, the *Play Mode* was fully developed and integrated in the Polariscope platform. Key implemented features included:

Predefined Viewing Modes:

- **Complete Story Mode:** Presents the entire narrative without filters, but in the segmented *Play Mode* format.
- **Summary Mode:** Shows a condensed version of the Story using only the most important contents as defined by the Story author.
- **Content-Type Mode:** Allows filtering the view by specific media types (e.g., images only, videos only).

User-Customized Mode: Configured through a step-by-step interactive questionnaire that captures personal preferences, including narrative pace, content density, preferred

media types, narration, and content popularity. Based on this configuration, a personalized viewing mode of the Story is created in the Play Mode.

Additional Features:

- Text-to-Speech (TTS) narration for enhanced immersion and accessibility.
- Creator-configured Summary Mode to support authors in highlighting key content.

This detailed implementation stage ensured that all viewing modes were ready for evaluation in laboratory tests (Phase 4).

3.4 Phase 4: Laboratory tests.

The functional *Play Mode* component was evaluated through structured Laboratory Tests with 10 participants. This non-probabilistic, convenience-based sample comprised university students from diverse academic backgrounds, ensuring continuity with the user profile tested in the focus group while expanding the sample size for more robust validation. The laboratory setting enabled greater control over some external variables and a systematic comparison of viewing modes, aligning the evaluation with the quasi-experimental principles of the study [18]. This dynamic facilitated data collection from participants. These tests included a combination of tasks, followed by a semi-structured interview with open questions and quantitative scales (Likert and SUS - System Usability Scale). Individual 40-minute sessions combined qualitative interviews with quantitative scales, for participants to compare Feed Mode and Play Mode, test predefined and personalised visualisation modes, and evaluate narration (TTS), immersion, usability, and the relevance of customisation. The main objectives were to validate the usability and usefulness of the features, measure user satisfaction, and identify possible improvements through structured questions and direct interaction with the platform.

3.5 Phase 5: Data analysis and results.

In the final phase, data from laboratory tests were analyzed using qualitative and quantitative approaches. Qualitative feedback from semi-structured interviews was thematically analyzed to identify recurring patterns, perceptions, and user expectations regarding the Play Mode and its viewing options. In parallel, quantitative data from Likert-scale responses and the System Usability Scale (SUS) were processed to assess usability, comprehension, immersion, satisfaction, and the perceived value of personalization in Story consumption. This enabled a robust evaluation of the *Play Mode*, reinforcing the validity of the findings and supporting the identification of the most valued personalized viewing solutions and their positive impact on the user experience.

4 Results

The presented evaluation outcomes focus on qualitative user preferences and their impacts on immersion, comprehension, and satisfaction. The results showed strong user preference for the Play Mode and personalized visualization features. Thus, the main conclusions regarding the study and *Play mode* were:

- Summary Mode was highly valued by the study participants, especially for cases like users with limited time on a daily basis.
- TTS narration increased immersion and accessibility of stories.
- Predefined modes helped novice users explore different visualization styles.
- The User-Customized Mode delivered the highest satisfaction and attraction levels.
- Users felt more appeal, agency and autonomy when consuming Stories.
- Users felt more engaged in the Polariscope platform because of the *Play Mode* and the customization featured in it.

5 Conclusion

This section synthesizes the main findings of the research and reflects on their implications for the design of personalized Digital Storytelling experiences. Based on the empirical evidence gathered throughout the different stages of the study, it outlines the contributions of the *Play Mode* to user experience, discusses its relevance within the broader Digital Storytelling landscape, and identifies directions for future research and development. The results gathered by this research allowed to understand that personalization in digital storytelling significantly enhances the user experience by improving immersion, comprehension, and engagement. The *Play Mode* features introduce a novel and effective approach to narrative visualization that aligns with trends in adaptive interfaces and personalization. Furthermore, it provides Polariscope with a competitive advantage while contributing meaningfully with insights to the broader Digital Storytelling research field. Future developments may include further user testing on the adaptive AI-driven personalization, the refinement of the text-to-speech feature with the implementation of additional voices and personalization options to the interface like the inclusion of adaptive visual features on the *Play Mode* interface like the ability to personalize colors, fonts and other graphical elements, and finally, further refinement of the User-Customized Mode questionnaire. Longitudinal studies could also assess sustained user engagement and long-term satisfaction trends. Additionally, targeting other groups, such as older users, and using in-app data collection methods for assessment during field trials will enable gathering more data and richer feedback from diverse types of users.

Acknowledgments. This work is financially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project 2022.04424.PTDC and the UIDB/05460/2020.

References

1. Roque, J., Almeida, P., Silva, T., Velhinho, A., Alves, M.: Personalized story-viewing experiences on the Polariscope platform. In: *Proc. ACM Int. Conf. Interactive Media Experiences Workshops (IMXw 2025)*, pp. 37–41. SBC, Porto Alegre (2025). <https://doi.org/10.5753/imxw.2025.3888>
2. Fatin, N., Rahman, N.A., Aboo Bakar, R.: Digital storytelling: A systematic review. *J. Engl. Lang. Teach. Innov. Mater.* **2**(2), 97–108 (2020). <https://doi.org/10.26418/jeltim.v2i2.42199>
3. Malita, L., Martin, C.: Digital storytelling as web passport to success in the 21st century. *Procedia Soc. Behav. Sci.* **2**, 3060–3064 (2010). <https://doi.org/10.1016/j.sbspro.2010.03.465>
4. Tian, Y., Suki, N.M., Omar, A.C.: Two decades of digital storytelling: A comprehensive bibliometric analysis. In: *Proc. IEEE Student Conf. Research and Development (SCORED 2021)*, pp. 125–129. IEEE (2021). <https://doi.org/10.1109/SCORED53546.2021.9652751>
5. Jenkins, H.: *Convergence Culture: Where Old and New Media Collide*. New York University Press, New York (2006).
6. Lévy, P.: *Cibercultura*. 1st edn. Editora 34, São Paulo (1997).
7. Durga, S., Mehrotra, V.: Communication and its vital role in human life. *Int. J. Human. Stud.* **6**(Suppl. 5) (2022). <https://doi.org/10.53730/ijhs.v6nS5.10005>
8. Sarica, H.Ç.: Emotions and digital storytelling in the educational context: A systematic review. *Rev. Educ.* **11**(3), e3430 (2023). <https://doi.org/10.1002/rev3.3430>
9. Rizvic, S., Djapo, N., Alispahic, F., Hadzihalilovic, B., Cengic, F.F., Imamovic, A., Okanovic, V., Boskovic, D.: Guidelines for interactive digital storytelling presentations of cultural heritage. In: *Proc. 9th Int. Conf. Virtual Worlds and Games for Serious Applications (VS-Games 2017)*, pp. 253–259. IEEE (2017). <https://doi.org/10.1109/VS-GAMES.2017.8056610>
10. Segel, E., Heer, J.: Narrative visualization: Telling stories with data. *IEEE Trans. Vis. Comput. Graph.* **16**(6), 1139–1148 (2010). <https://doi.org/10.1109/TVCG.2010.179>
11. He, X.: Interactive mode of visual communication based on information visualization theory. *Comput. Intell. Neurosci.* **2022**, Article ID 4482669 (2022). <https://doi.org/10.1155/2022/4482669>
12. Pardo, A.: One dataset – three stories: Data storytelling for climate change awareness. In: *Proc. Int. Conf. Information Visualisation*, pp. 194–197. IEEE (2023). <https://doi.org/10.1109/IV60283.2023.00042>
13. Creswell, J.W., Plano Clark, V.L.: *Designing and Conducting Mixed Methods Research*. 3rd edn. Sage, Thousand Oaks (2017).
14. Ivankova, N.V., Creswell, J.W., Stick, S.L.: Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods* **18**(1), 3–20 (2006). <https://doi.org/10.1177/1525822X05282260>
15. Wang, F., Hannafin, M.J.: Design-based research and technology-enhanced learning environments. *Educ. Technol. Res. Dev.* **53**(4), 5–23 (2005).
16. Vaezi, H., Moonaghi, H.K., Golbaf, R.: Design-based research: Definition, characteristics, application and challenges. *J. Educ. Black Sea Reg.* **5**(1), 26–35 (2019). <https://doi.org/10.31578/jeds.v5i1.185>
17. Almutrafi, F.: The what, why and how of conducting focus-group research. *Int. J. Lang. Linguist.* **7**(5), 235 (2019). <https://doi.org/10.11648/j.ijll.20190705.18>
18. White, H., Sabarwal, S.: *Quasi-experimental Design and Methods*. UNICEF Office of Research, Florence (2014).

jauti2025.web.ua.pt

28
NOVEMBER

UNIVERSIDADE DE AVEIRO
Portugal



JAUTI'25