



JORNADAS IBEROAMERICANAS DE INTERACCIÓN HUMANO - COMPUTADORA 2025

Interacción Humano-Tecnología
en la Era de la Inteligencia Artificial
2



AUTORIDADES UAZ

Dr. Rubén de Jesús Ibarra Reyes

Rector

Dr. Ángel Román Gutiérrez

Secretario General

Dr. Hans Hiram Pacheco García

Secretario Académico

Dr. Carlos Bautista Capetillo

Coordinador de Investigación y Posgrado

M.I. María de Lourdes Oliván Tiscareño

Coordinadora del Área de Ingenierías y Tecnología

Dr. Jorge Issac Galván Tejada

Director de la Unidad Académica de Ingeniería Eléctrica

AUTORIDADES COZCYT

Dr. Hamurabi Gamboa Rosales

Director del Consejo Zacatecano de Ciencia, Tecnología e Innovación



Diseño Editorial: Programa Editorial UAZ

Maquetación y portada: Huizilopoztli Luna García, Jaime Mercado Reyna

Portada: Jaime Mercado Reyna

Primera Edición: 2025

Universidad Autónoma de Zacatecas

"Francisco García Salinas"

Centro de Información Siglo XXI, 3er piso,

Campus UAZ Siglo XXI,

Carretera Zacatecas-Guadalajara,

Km 6, Col. Ejido Escondida,

C.P. 98000, Zacatecas, Zac.

programaeditorial@uaz.edu.mx

Archivo Digital: descarga online

ISBN 978-607-555-301-6

Se prohíbe la reproducción total o parcial de esta obra, por cualquier modo electrónico o mecánico, sin la autorización de la institución editora.

El contenido de esta obra es responsabilidad de cada autor.



COMITÉ EDITORIAL

Director Editorial
Coordinador Editorial

Huizilopoztli Luna García

Compiladores

Huizilopoztli Luna García
Cesar A. Collazos
José Ma. Celaya Padilla

Editores Literarios

Huizilopoztli Luna García
Luis Alberto Flores Chaires
Jaime Mercado Reyna
Jorge Alejandro Morgan Benita
Carlos Humberto Espino Salinas
Sandra Elizabeth Flores

Diseñador

Jaime Mercado Reyna

Autores

María Guadalupe Celaya Padilla, et al.
Karla M. García-Castillo, et al.
Isaías R. Valdivia, et al.
Gabriel Alfonso Rodríguez Puerto, et al.
Vinicius Alves Marques, et al.
Felipe Chuang, et al.
Carlos A. Meneses Escobar, et al.
Eduar Rodolfo Becerra-Martínez, et al.
Gabriel Batista Cristiano, et al.
José G. González Rodarte, et al.
Juan Fernando Florez M., et al.
Cristina Collado Piñana, et al.
Ana Covadonga Schleissner Granda, et al.
Diany Isabel García-Carrascal, et al.
Cristian Plazas, et al.
Juan Sebastián Carballo Ñustes, et al.



ACTAS DE LAS JORNADAS

Las memorias de las 11 Jornadas Iberoamericanas de Interacción Humano-Computadora (11JIHC'25) reúnen una selección de los trabajos aceptados y presentados en modalidad de artículos cortos (*short papers*) y artículos largos (*full papers*) durante el evento realizado en la ciudad de Zacatecas, México, del 7 al 9 de mayo de 2025, bajo un formato híbrido. Las Jornadas fueron organizadas por la Universidad Autónoma de Zacatecas "Francisco García Salinas", en colaboración con el Consejo Zacatecano de Ciencia, Tecnología e Innovación (COZCYT) y la Red Iberoamericana HCI-COLLAB.

El eje temático "Interacción Humano-Tecnología en la Era de la Inteligencia Artificial" orientó esta edición hacia el análisis crítico de cómo las tecnologías emergentes transforman las experiencias humanas y sociales, impulsando la accesibilidad, la inclusión y la colaboración en sectores clave como la educación, la salud, la industria y el entretenimiento. En este marco, los artículos enviados proceden de los talleres *7° Workshop en Sistemas Interactivos Inteligentes*, *4° Workshop sobre Industria 4.0 e Interacción Humano-Computadora*, *4° Workshop de Trabajo Colaborativo y Aprendizaje Colaborativo*, *3^{er} Workshop en Cognición e Interacción* y *2° Workshop sobre Accesibilidad e Inclusión*, los cuales aportan perspectivas metodológicas y aplicaciones avanzadas que fortalecen el entendimiento interdisciplinar de la Interacción Humano-Computadora en escenarios mediados por inteligencia artificial.

En atención a los lineamientos editoriales, los trabajos distinguidos como mejores contribuciones no se incluyen en estas memorias, dado que serán publicados en la serie *Communications in Computer and Information Science (CCIS)* de Springer.

Mayo 2025



ORGANIZACIÓN

AUTORIDADES DE LAS JORNADAS

Dr. Rubén de Jesús Ibarra Reyes	Rector
Dr. Ángel Román Gutiérrez	Secretario General
Dr. Hans Hiram Pacheco García	Secretario Académico
Dr. Carlos Bautista Capetillo	Coordinador de Investigación y Posgrado
M.I. María de Lourdes Oliván Tiscareño	Coordinadora del Área de Ingenierías y Tecnología
Dr. Jorge Issac Galván Tejada	Director de la Unidad Académica de Ingeniería Eléctrica
Dr. Hamurabi Gamboa Rosales	Director del Consejo Zacatecano de Ciencia, Tecnología e Innovación

COORDINADORES GENERALES

Huizilopoztli Luna García	Universidad Autónoma de Zacatecas, México
César Alberto Collazos Ordóñez	Universidad del Cauca, Colombia

COMITÉ ORGANIZADOR

Luis Alberto Flores Chaires	Universidad Autónoma de Zacatecas, México
José María Celaya Padilla	Universidad Autónoma de Zacatecas, México
Jorge Issac Galván Tejada	Universidad Autónoma de Zacatecas, México
Hamurabi Gamboa Rosales	Consejo Zacatecano de Ciencia, Tecnología e Innovación (COZCYT), México
Toni Granollers Saltiveri	Universitat de Lleida, España
Jaime Muñoz Arteaga	Universidad Autónoma de Aguascalientes, México



COMITÉ LOCAL

Ana Gabriela Sánchez Reyna	Instituto Tecnológico de Aguascalientes, México
Arturo Plascencia Rodríguez	Universidad Autónoma de Zacatecas, México
Carlos H. Espino Salinas	Universidad Autónoma de Zacatecas, México
Esmeralda Carrillo Chávez	Universidad Autónoma de Zacatecas, México
Holman S. Cabezas Mondragón	Universidad Autónoma de Zacatecas, México
Jaime Mercado Reyna	Instituto Tecnológico de Aguascalientes, México
Jorge Alejandro Morgan Benita	Universidad Autónoma de Zacatecas, México
José Germán González Rodarte	Universidad Autónoma de Zacatecas, México
Oscar Armando Veruete Macías	Universidad Autónoma de Zacatecas, México
Rubén de Jesús Rivera Medina	Universidad Autónoma de Zacatecas, México

COMITÉ DE PROGRAMA

Alicia Mon	Universidad Nacional de La Matanza, Argentina
Andrés Rodríguez	Universidad Nacional de La Matanza, Argentina
Christian Sturm	Universidad Tecnológica de Ingolstadt, Alemania
Cibelle Albuquerque de la Higuera Amato	Universidade Presbiteriana Mackenzie, Brasil
Francisco Luis Gutiérrez Vela	Universidad de Granada, España
Horacio René Del Giorgio	Universidad Nacional de La Matanza, Argentina
Ismar Frango Silveira	Universidade Presbiteriana Mackenzie, Brasil
Jeferson Arango López	Universidad de Caldas, Colombia
Maria Amelia Eliseo	Universidade Presbiteriana Mackenzie, Brasil
Nancy Estévez Pérez	Centro de Neurociencias de Cuba, Cuba
Omar Correa Madrigal	Universidad de las Ciencias Informáticas
Pablo H. Ruiz	Corporación Universitaria Comfacauca, Colombia
Patricia Paderewski Rodríguez	Universidad de Granada, España
Susana Bautista Blasco	Universidad Francisco de Vitoria, España



Valéria Farinazzo Martins
Vanessa Agredo Delgado
Wilson J. Sarmiento

Universidade Presbiteriana Mackenzie, Brasil
Corporación Universitaria Comfacaucua, Colombia
Universidad Militar Nueva Granada, Colombia

COMITÉ CIENTÍFICO

Christian Sturm
Patricia Appel
Silvana Vanesa Aciar

Universidad Tecnológica de Ingolstadt, Alemania
Universidad Tecnológica de Ingolstadt, Alemania
Instituto de Ciencias Astronómicas de la Tierra y el Espacio,
Argentina

Alicia Mon
Horacio René Del Giorgio
Andrés Rodríguez
Cecilia Challiol
Diego Torres
Gustavo Rossi
Julián Grigera
Luis Mariano Bibbo
Germán Ezequiel Lescano
Rosanna Costaguta
Carolina Saavedra Sueldo

Universidad Nacional de La Matanza, Argentina
Universidad Nacional de La Matanza, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de la Plata, Argentina
Universidad Nacional de Santiago del Estero, Argentina
Universidad Nacional de Santiago del Estero, Argentina
Universidad Nacional del Centro de la Provincia de Buenos
Aires, Argentina

Geraldina Roark

Universidad Nacional del Centro de la Provincia de Buenos
Aires, Argentina

Ivo Perez Colo

Universidad Nacional del Centro de la Provincia de Buenos
Aires, Argentina

María Fernanda Montoya Vega
Marcelo Morandini
André Constantino da Silva
Daniela Marques

Universidad de Monash, Australia
Universidad de São Paulo, Brasil
Instituto Federal de São Paulo, Brasil
Instituto Federal de São Paulo, Brasil



Soares de Oliveira Neto João	Universidade Federal de Bahia, Brasil
Marcelo de Paiva Guimarães	Universidade Federal de São Paulo, Brasil
Alexandre Cardoso	Universidade Federal de Uberlândia
André Luiz Satoshi Kawamoto	Universidade Tecnológica Federal de Paraná, Brasil
María Amelia Eliseo	Universidade Presbiteriana Mackenzie
Cibelle Albuquerque de la Higuera Amato	Universidade Presbiteriana Mackenzie
Ismar Frango	Universidade Presbiteriana Mackenzie
Valeria Farinazzo Martins	Universidade Presbiteriana Mackenzie
Nicolás Esteban Matus Palma	Pontificia Universidad Católica de Valparaíso, Chile
Sergio Ochoa	Universidad de Chile
Jaime Díaz Arancibia	Universidad de la Frontera, Chile
Maribel Ramos Hernández	Universidad de Playa Ancha, Chile
Claudio Alfredo Rubio Naranjo	Universidad de Talca, Chile
Felipe Andrés Besoain Pino	Universidad de Talca, Chile
Andrés Felipe Solis Pino	Corporación Universitaria Comfacaucá, Colombia
Vanessa Agredo	Corporación Universitaria Comfacaucá, Colombia
Luis Freddy Muñoz Sanabria	Fundación Universitaria de Popayán, Colombia
Leandro Flórez Aristizábal	Institución Universitaria Antonio José Camacho, Colombia
Manuel Alejandro Pastrana Pardo	Institución Universitaria Antonio José Camacho, Colombia
Beatriz Eugenia Marin Ospina	Institución Universitaria Antonio José Camacho, Colombia
Marta Cecilia Camacho Ojeda	Institución Universitaria Colegio Mayor del Cauca, Colombia
Andrés Solano	Universidad Autónoma de Occidente, Colombia
Javier A. Jiménez Toledo	Universidad CESMAG, Colombia
Jeferson Arango López	Universidad de Caldas, Colombia
Oscar Revelo Sánchez	Universidad de Nariño, Colombia
Manuel E. Bolaños González	Universidad de Nariño, Colombia
Anyela P. Villamizar Carrillo	Universidad de Pamplona, Colombia
Victor Manuel Peñeñory	Universidad de San Buenaventura Cali, Colombia



Cesar A. Collazos	Universidad del Cauca, Colombia
Pablo Ruiz	Universidad del Cauca, Colombia
Maria Lili Villegas Ramírez	Universidad del Quindío, Colombia
William Joseph Giraldo Orozco	Universidad del Quindío, Colombia
Yesenia Areniz Arévalo	Universidad Francisco de Paula Santander Ocaña, Colombia
Luis Anderson Coronel Rojas	Universidad Francisco de Paula Santander Ocaña, Colombia
Eduar Bayona Ibáñez	Universidad Francisco de Paula Santander Ocaña, Colombia
Dewar Rico Bautista	Universidad Francisco de Paula Santander Ocaña, Colombia
Beatriz E. Rodríguez Chinchilla	Universidad Francisco de Paula Santander Ocaña, Colombia
Alexander Guerrero Avendaño	Universidad Francisco de Paula Santander Ocaña, Colombia
Wilson J. Sarmiento Manrique	Universidad Militar Nueva Granada, Colombia
Byron Pérez Gutiérrez	Universidad Militar Nueva Granada, Colombia
Carlos Javier Latorre Rojas	Universidad Militar Nueva Granada, Colombia
Gustavo E. Constain Moreno	Universidad Nacional Abierta y a Distancia, Colombia
María Consuelo Rodríguez	Universidad Nacional Abierta y a Distancia, Colombia
José Miguel Herrán Suárez	Universidad Nacional Abierta y a Distancia, Colombia
Diana Castañeda	Universidad Nacional Abierta y a Distancia, Colombia
Marco Javier Suarez Baron	Universidad Pedagógica y Tecnológica de Colombia
Raúl Eduardo Rodríguez Ibáñez	Universidad Simón Bolívar, Colombia
Jessica Lorena Leal Pabón	Universidad Simón Bolívar, Colombia
Cesar Jaramillo	Universidad Tecnológica de Pereira, Colombia
Julio César Chavarro Porras	Universidad Tecnológica de Pereira, Colombia
Oscar Alberto Henao Gallo	Universidad Tecnológica de Pereira, Colombia
Carlos A. Meneses Escobar	Universidad Tecnológica de Pereira, Colombia
Jairo Alberto Villegas Flórez	Universidad Tecnológica de Pereira, Colombia
Fulvio Lizano Madriz	Universidad Nacional de Costa Rica, Costa Rica
Omar Correa Madrigal	Universidad de las Ciencias Informáticas, Cuba
Otto Parra González	Universidad de Cuenca, Ecuador
Habiib Fardoun Canadian	Universidad de Dubai, Emiratos Árabes Unidos
José Antonio Macías Iglesias	Universidad Autónoma de Madrid, España



Yussy Chinchay Manco	Universidad Autónoma de Madrid, España
Ruth S. Contreras Espinosa	Universidad de Barcelona, España
Elena Navarro	Universidad de Castilla - La Mancha, España
Juan Enrique Garrido	Universidad de Castilla - La Mancha, España
Manuel Ortega Cantero	Universidad de Castilla - La Mancha, España
María Lozano	Universidad de Castilla - La Mancha, España
Miguel Ángel Redondo Duque	Universidad de Castilla - La Mancha, España
Víctor M. R. Penichet	Universidad de Castilla - La Mancha, España
Jose García Alonso	Universidad de Extremadura, España
Francisco Gutiérrez Vela	Universidad de Granada, España
Patricia Paderewski Rodríguez	Universidad de Granada, España
Ramon Mas Sansó	Universidad de las Islas Baleares, España
Alicia García Holgado	Universidad de Salamanca, España
Francisco José García Peñalvo	Universidad de Salamanca, España
Eva Cerezo	Universidad de Zaragoza, España
Jesús Gallardo Casero	Universidad de Zaragoza, España
Sandra Baldassarri	Universidad de Zaragoza, España
Sergio Albiol Perez	Universidad de Zaragoza, España
Susana Bautista Blasco	Universidad Francisco de Vitoria, España
Natalia Padilla Zea	Universidad Internacional de la Rioja, España
Rosa Maria Gil Iranzo	Universidad de Lleida, España
Toni Granollers	Universidad de Lleida, España
David A. Céspedes Hernández	Octahedroid Systems, Estados Unidos de América
Alexandre Gomes de Siqueira	Universidad de Florida, Estados Unidos de América
John Alexis Guerra Gomez	Universidad del Nordeste, Estados Unidos de América
Pradeep Kumar Sharma	ServiceNow, Estados Unidos de América
bhinav Mehta Aditya Gupta	Amazon.com, Estados Unidos de América
Marco Winckler	Université Côte d'Azur, Francia
Palanque Phil	ICS-IRIT, University Toulouse III, Francia
Huizilopoztli Luna García	Universidad Autónoma de Zacatecas, México



Luis Alberto Flores Chaires	Universidad Autónoma de Zacatecas, México
Carlos Eric Galván Tejada	Universidad Autónoma de Zacatecas, México
Carlos Humberto Espino Salinas	Universidad Autónoma de Zacatecas, México
Jorge Alejandro Morgan Benita	Universidad Autónoma de Zacatecas, México
José Guadalupe Arceo Olague	Universidad Autónoma de Zacatecas, México
Roberto Solís Robles	Universidad Autónoma de Zacatecas, México
José María Celaya Padilla	Universidad Autónoma de Zacatecas, México
Héctor Cardona Reyes	Centro de Investigación en Matemáticas, México
Ana Gabriela Sánchez Reyna	Instituto Tecnológico de Aguascalientes, México
Carlos Alberto Domínguez Báez	Instituto Tecnológico de Aguascalientes, México
Ricardo Mendoza González	Instituto Tecnológico de Aguascalientes, México
Raúl Antonio Aguilar Vera	Universidad Autónoma de Yucatán, México
Alejandro Moreno Cruz	Universidad Autónoma de Aguascalientes, México
Jaime Muñoz Arteaga	Universidad Autónoma de Aguascalientes, México
Julio César Enríquez González	Universidad Autónoma de Aguascalientes, México
Alfredo García Suárez	Benemérita Universidad Autónoma de Puebla, México
Guillermina Sánchez Román	Benemérita Universidad Autónoma de Puebla, México
Josefina Guerrero García	Benemérita Universidad Autónoma de Puebla, México
Juan Manuel González Calleros	Benemérita Universidad Autónoma de Puebla, México
Claudia Blanca González Calelleros	Universidad Veracruzana, México
Gisela Torres de Clunie	Universidad Tecnológica de Panamá, Panamá
Arturo Moquillaza	Pontificia Universidad Católica del Perú, Perú
Fiorella Falconi Trauco	Pontificia Universidad Católica del Perú, Perú
Klinge O. Villalba Condori	Universidad Católica de Santa María, Perú
Franci Suni Lopez	Universidad de Lima, Perú
Fernando Moreira	Universidade Portucalense, Portugal
Maria João Ferreira	Universidade Portucalense, Portugal
Antonio Silva Sprock	Universidad Central de Venezuela, Venezuela



ÍNDICE

7º WORKSHOP SOBRE SISTEMAS INTERACTIVOS INTELIGENTES

A Prototype Interface for Emotion Visualization in Climate Change Discourse <i>María Guadalupe Celaya Padilla, Huizilopoztli Luna-García, Roberto Solís-Robles, Carlos H. Espino-Salinas, Ana G. Sánchez-Reyna, José G. Arceo-Olague.....</i>	1-10
Aplicación Móvil con Realidad Aumentada para el Centro Histórico de Zacatecas <i>Karla M. García-Castillo, Huizilopoztli Luna-García, Roberto Solís-Robles, Hamurabi Gamboa-Rosales, Cristian Barría-Huidobro, Klinge O. Villalba-Condori.....</i>	11-20
Diseño de una Aplicación Educativa en Realidad Virtual para Fomentar Habilidades Críticas en el Pensamiento Computacional <i>Isaías R. Valdivia, Huizilopoztli Luna-García, José M. Celaya-Padilla, Wilson J. Sarmiento, Carlos H. Espino-Salinas, Klinge O. Villalba-Condori.....</i>	21-27
Evaluating the Impact of Headset Weight on User Comfort in Immersive Visualization <i>Gabriel Alfonso Rodríguez Puerto, Nicolás Daniel García Sierra, Daryi Valentina Vargas Umbarila, Carlos J. Latorre-Rojas, Wilson J. Sarmiento.....</i>	28-35
Sistema Inteligente de Atendimento a Emergências: Viabilizando Gêmeo Digital e Otimização de Recursos em Cidades Inteligentes <i>Vinicius Alves Marques, Leandro Pupo Natale.....</i>	36-43

4º WORKSHOP SOBRE INDUSTRIA 4.0 Y HCI

Challenges and Strategies for Industry 4.0 Adoption in the Brazilian Wholesale Sector <i>Felipe Chuang, Maria Amelia Eliseo, Daniela Vieira Cunha.....</i>	44-55
Detección del Grado de Madurez en Aguacate Hass Usando Procesamiento Digital de Imágenes <i>Carlos Augusto Meneses Escobar, Julio Cesar Chavarro Porras, Cesar Augusto Jaramillo Acevedo, Julio Cesar López Betancu.....</i>	56-64
Emerging Technologies, HCI, and Agile Methodologies in Teacher Professional Development: A Study with 3D Printing and eduScrum <i>Eduar Rodolfo Becerra-Martínez, Juan Manuel González-Calleros, Erika Annabel Martínez Mirón.....</i>	65-75



Recommendations Guide for Adopting Industry 4.0 Practices for Small and Medium-Sized Brazilian Companies in the Food Sector

Gabriel Batista Cristiano, Fabio Tanikawa, Huizilopoztli Luna-García, Maria Amelia Eliseo.....

76-86

4^o WORKSHOP SOBRE TRABAJO COLABORATIVO Y APRENDIZAJE COLABORATIVO

Prototipo de Aplicación Móvil para la Detección de Tendencias Auto-Lesivas en Estudiantes de Medicina

José G. González Rodarte, Esmeralda Carrillo Chávez, Jaime Mercado Reyna, Huizilopoztli Luna-García.....

87-99

3^{ER} WORKSHOP SOBRE COGNICIÓN E INTERACCIÓN

Aula Virtual Inmersiva de Señalización Deíctica Grupal para Educación Remota

Juan Fernando Florez M., Diego Fernando Valencia M., Elena Muñoz E., Cesar A. Collazos, Julio A. Hurtado, Christian Sturm.....

100-107

2^o WORKSHOP SOBRE ACCESIBILIDAD E INCLUSIÓN

Day a Day Communicator: Aplicación Web Destinada a Mejorar la Comunicación Asistida de Personas con Trastorno del Espectro Autista

Cristina Collado Piñana, Susana Bautista Blasco.....

108-113

KetoGlut Planner: Aplicación de Ayuda para Personas con Deficiencia de Glut-1.

Ana Covadonga Schleissner Granda, Susana Bautista Blasco.....

114-121

Multiplatform System for the Prediagnosis of Hearing Loss: Implementation and Validation.

Diany Isabel García-Carrascal, Jose Leonardo Quintero-Leon, Magreth Rossio Sanguino-Reyes, Huizilopoztli Luna-García, Dewar Rico-Bautista.....

122-133

Narrativa Semántica y Características Prosódicas para Audiodescripción

Cristian Plazas, Juan Sebastian Cardenas, Christian Quintero.....

134-143

Síntesis de Voz Emocional Basada en Prosodia para Audiodescripción

Juan Sebastián Carballo Nustes, Alexander Roza-Torres, Christian Quintero.....

144-155



7º WORKSHOP SOBRE SISTEMAS INTERACTIVOS INTELIGENTES

A Prototype Interface for Emotion Visualization in Climate Change Discourse

María Guadalupe Celaya Padilla¹, Huizilopoztli Luna-García¹, Roberto Solís-Robles¹,
Carlos H. Espino-Salinas¹, Ana G. Sánchez-Reyna², José G. Arceo-Olague¹
¹ Universidad Autónoma de Zacatecas, México
34152203@uaz.edu.mx, hlugar@uaz.edu.mx, rsolis@uaz.edu.mx,
carlospino@uaz.edu.mx, arceojg@uaz.edu.mx
² Tecnológico Nacional de México, Instituto Tecnológico de Aguascalientes, México
agsreyna@uaz.edu.mx

Abstract. Climate change evokes strong emotions that influence perception and social action. This study presents an interface for analyzing and visualizing emotions in online forums such as Reddit or Twitter, using natural language processing techniques. A heuristic and user-centered evaluation was conducted, revealing that the interface is perceived as intuitive and clear, although with opportunities for improvement in visualization customization. These results highlight its potential for researchers, activists, and the public. The tool facilitates the interpretation of emotional trends, promoting a better understanding of the emotional impact of climate change.

Keywords: Sentiment Analysis, Climate Change, Online Forums, NLP, Emotion Detection, Human-Data Interaction, Social Media Analysis.

1 Introduction

Climate change is one of the most significant challenges for our generation, with an impact that goes beyond environmental concerns, affecting the emotions and perceptions of the general public. Capturing and understanding these emotions is essential for developing effective strategies that promote positive actions and constructive dialogue. In this regard, online forums have become important spaces where people express their thoughts and emotions [4]. However, manually analyzing large volumes of emotional data can be tedious and inefficient.

Current literature suggests the use of digital tools to address this challenge. For example, *Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data* [4] highlights how digital platforms capture predominant emotions that can guide communication strategies useful in areas such as marketing. Likewise, *The Evolution of Sentiment Analysis* [1] explores modern techniques that facilitate the processing of emotions in large volumes of text.

Despite advancements in emotional analysis tools, a gap remains in how these results can be translated into comprehensible platforms for the public. This work seeks to bridge that gap through the design of an intuitive interface that not only analyzes emotions but also fosters collective understanding of climate change. The platform was designed with a user-centered approach, prioritizing the needs and expectations of end users, ensuring that the system is functional, intuitive, and relevant.

Main Contributions. This paper presents:

- A user-centered interface for visualizing emotional trends in climate change forums.
- A lightweight prototype tested with users based on ISO/IEC 29110.
- A discussion on the challenges of translating NLP results for non-technical audiences.

2 Literature review

In the publication *Emotions and the Dissemination of Climate Change News on Social Media: Influence of Habits, Prior Attitudes, and Uses and Gratifications Among University Students* [12], the main objective is to analyze how emotions influence university students' decisions to share climate change news on social media. This study identifies which emotions are most relevant in this process and explores news consumption habits and prior attitudes toward climate change, as well as how they moderate this emotional influence. However, its approach does not directly address the emotions expressed in posts or the emotional response of users after sharing such news. Instead, it focuses on the perspective prior to the act of sharing content, distinguishing it from this work.

On the other hand, the *ClimateNLP* workshop applies natural language processing techniques to analyze climate change discourse, focusing on quantifying sentiment in posts related to this topic [3]. This workshop stands out for its approach and contribution to data analysis but lacks a platform that translates these results into accessible tools for the public.

The thesis *Sentiment Analysis in Social Media* presented by Rafael Reyes Podadera [10] examines the evolution of verbal communication through natural language in digital platforms. It employs artificial intelligence techniques such as Naive Bayes, support vector machines, VADER, and TextBlob to analyze sentiments expressed on platforms like Twitter. Although this work has a similar focus, it is oriented toward a business perspective, aimed at evaluating public perception of brands on social media.

Despite the diversity of related studies, many of them are limited to technical and academic analysis without presenting platforms for the public. This contrasts with the proposed interface, which seeks not only to evaluate sentiment classification tools in the context of climate change but also to translate the results into an informative visualization designed for educational campaigns or environmental awareness initiatives.

3 Methodology

To ensure the usability and relevance of the interface, a user-centered design methodology was adopted. This approach allowed for the identification of issues, exploration of solutions, and validation of functionalities throughout the development process. Additionally, a lightweight sentiment analysis model (VADER) was integrated into the system to process real Reddit data. This component allowed the prototype to demonstrate basic classification and visualization capabilities, enabling initial testing with real-world content (see Fig. 1 as an example).

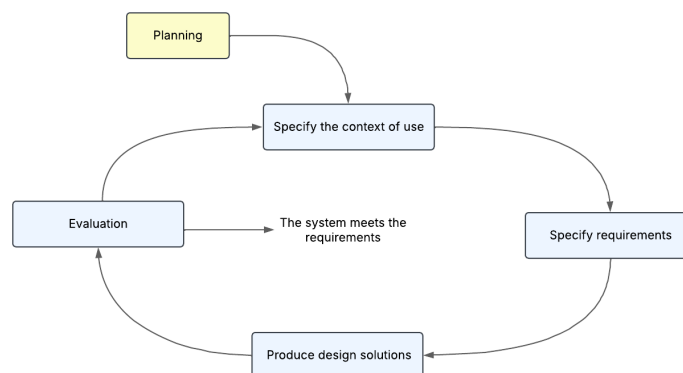


Fig. 1. UCD model. (Source: Own elaboration).

The development of this prototype followed principles aligned with the ISO/IEC 29110 standard, designed for small entities. This approach allowed for the structured organization and management of the project, starting with proper requirement gathering through focus groups and interviews, followed by iterative design and validation. Additionally, incorporating feedback from end users ensured that the system was functional, intuitive, and relevant, meeting the initially defined objectives [8].

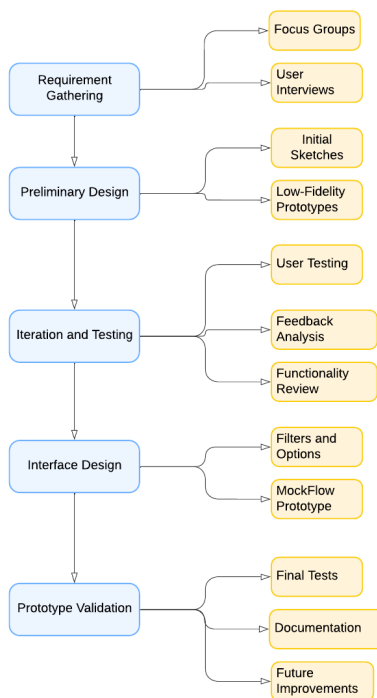


Fig. 2. Prototyping process structure. (Source: Own elaboration).

3.1 Requirement gathering

In this initial stage, activities were conducted to identify user needs and establish the foundation for developing an intuitive and useful interface.

Focus Group Organization

A discussion session was conducted with university students selected for their knowledge in design and emotion analysis [7].

- Activities included:
 - **Usage Identification:** Participants discussed how they envision an ideal tool for emotion analysis in topics such as climate change [2].
 - **Project Challenges:** Possible issues in interpreting emotional data were examined, such as understanding graphs or accessibility of information [15].
 - **Key Features:** Students highlighted essential functionalities they considered necessary (clear visualizations, customizable filters, etc.) [7].

Interviews

In addition to focus groups, interviews were conducted to gain a more detailed perspective on user needs [6].

The interviews were directed at:

- **Individuals over 18 years old:** Including fields such as engineering and social sciences, aiming to cover a variety of perspectives and technical needs to understand how people might interact with the tool.

During the interviews, the following topics were addressed:

- Current tools they use or know for emotion analysis.
- Their preferences for data presentation (graphs, tables, textual explanations).
- Technological barriers, such as familiarity with analysis systems or accessibility issues.

The responses were organized and served as input for the preliminary design of the interface.

4 Interface design

The interface design focuses on analyzing emotions in online forums, identifying positive, negative, and neutral sentiments, as well as tracking trends over time while ensuring data clarity and accessibility [13]. The following key elements were implemented:

4.1 Core features

- **Emotion Visualization:** The interface displays emotional trends derived from simplified sentiment classification, using rule-based models such as VADER. This enables the generation of illustrative visualizations based on text content from Reddit related to climate change. Although preliminary, the integration demonstrates how emotional insights could be presented to end users through an accessible UI.
- **Visualization:** Graphs display detected emotions and trends over time [14]. Figures 5 and 6 use mock data to simulate how results would be displayed, due to limitations in deploying a full backend for this prototype version
- **Result Exporting:** Generates reports in formats such as PDF and CSV [9].
- **Social Media Sharing:** Allows users to share results on platforms like X, Facebook, and WhatsApp.
- **Custom Filters:** Enables filtering of specific information for detailed analysis.

4.2 User interface

The interface is designed to be intuitive and user-friendly, ensuring easy navigation and accessibility to all features, even for users without technical knowledge [13].

Table 1. Summary of system considerations, requirements, and features.

Section	Details
Technical Considerations	Compatible with Chromium-based browsers. Requires a stable internet connection. Does not store user data to ensure privacy.
System Requirements	Hardware: 8-core processor. 16 GB of RAM. 500 GB SSD storage. High-end GPU for graphical visualizations. Software: Developed in Python. Compatible with modern browsers (Chrome, Edge).

User Characteristics	Researchers with experience in data analysis. Activists with basic knowledge of technology. General public interested in emotion analysis.
Functional Requirements	Filter characteristics such as gender and aggressiveness. Select topics to analyze. Generate emotion visualizations (radar and pie charts). Export data in PDF and CSV formats. Share results on social media. Display word clouds.
Non-functional Requirements	Intuitive and easy-to-use interface. Security and privacy of processed data. Efficient performance for processing large volumes of text. High availability and compatibility with popular browsers. Scalability to support increases in user load.

The following sequence diagram illustrates the interaction flow for data export within the system, highlighting the interactions between the user and the main modules. The process begins with access to the interface, where the user selects a topic by applying filters. The system processes the information through the model and generates visualizations with the processed data.

When the export option is selected, the system converts the information into a PDF or CSV file, ensuring compatibility across different browsers and data protection. Validations are included to verify the correct generation of the file and to manage potential errors, notifying the user of the process outcome.

This workflow ensures an efficient, intuitive, and secure experience in managing and exporting information.

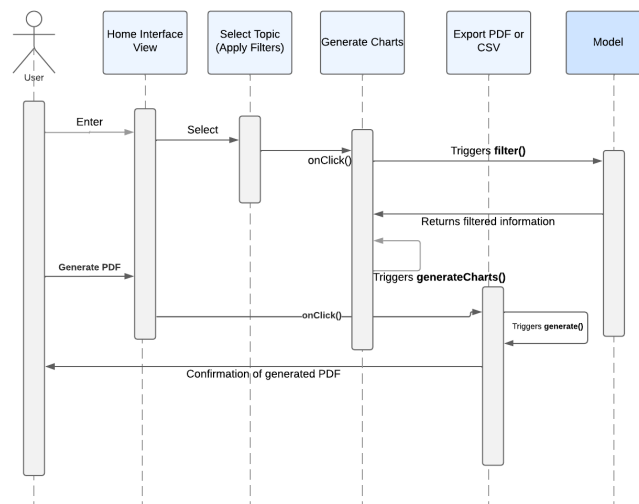


Fig. 3. System sequence diagram for generating CSV or PDF. (Source: Own elaboration).

4.3 System attributes

The system is designed to be maintainable and scalable, allowing for the future addition of new functionalities, such as more in-depth analysis or support for additional languages.

5 Results

Based on the expectations identified during the focus groups and interviews, the interface is expected to achieve the following objectives:

- Provide an easy-to-use and accessible tool, even for users with little experience in data analysis.
- Offer clear visualizations that allow for quick and effective interpretation of emotions.
- Facilitate the export of results in formats useful for research, presentations, and communication strategies.

5.1. Interface

This section presents screenshots of the developed interface for emotion analysis.

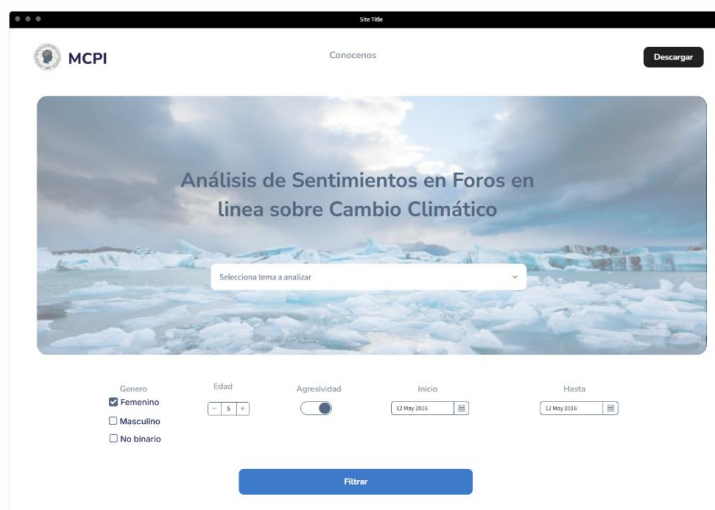


Fig. 4. General view of the interface.

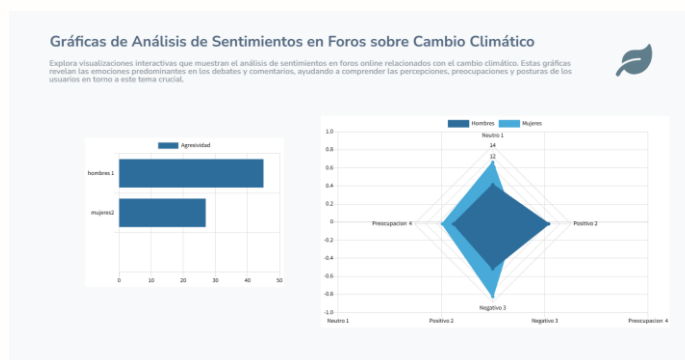


Fig. 5. Detailed view of emotion analysis.



Fig. 6. Word cloud and pop-up window for sharing.



Fig. 7. Mobile view upon entering the system.

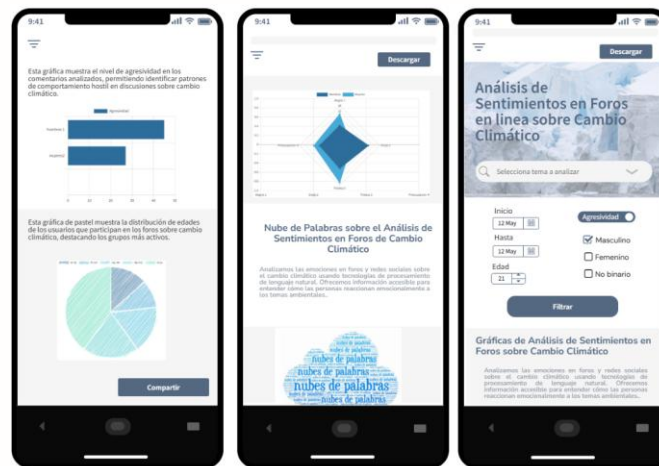


Fig. 8. Mobile view when applying filters.

Although the user-centered methodology allowed for the development of a functional and relevant tool, there are important limitations:

- The test participants were primarily university students, which may limit the generalization of the findings.
- Communities with limited access to technology were not included, which could affect the system’s usability in such contexts.

6 Evaluation

To assess the system's usability, a heuristic evaluation was conducted alongside semi-structured interviews with users over 18 years old. The evaluation focused on three key aspects: ease of use, visual design, and overall satisfaction.

6.1 Methodology

The evaluation was carried out in two stages:

- **Heuristic Evaluation:** Usability principles based on Nielsen's heuristics [5] were applied to identify design and navigation issues.
- **User Interviews:** Interviews were conducted with five participants, who evaluated the system while performing specific tasks such as generating graphical visualizations and exporting data [6].

During the session, participants were asked to describe their experience and rate the interface on a scale from 1 to 5 in the following aspects:

1. Clarity of the graphics.
2. Ease of navigation.
3. Intuitiveness of the interface.
4. Overall satisfaction with the tool.

6.2 Results

The evaluation results are presented in **Table 2**. Most participants found that the clarity of the graphics was a strong point, while the generation of visualizations with different approaches was identified as an area for improvement.

Table 2. Usability evaluation results.

Evaluated Aspect	Mean (1-5)	Standard Deviation
Graph Clarity	4.4	0.4898
Navigation Ease	4.2	0.7483
Interface Intuitiveness	4.0	0.6324
Overall Satisfaction	4.4	0.4898

Additionally, Figure 9 presents a comparative visualization of the evaluations. It can be observed that clarity of graphics received the highest rating, while interface intuitiveness showed greater variability in responses.

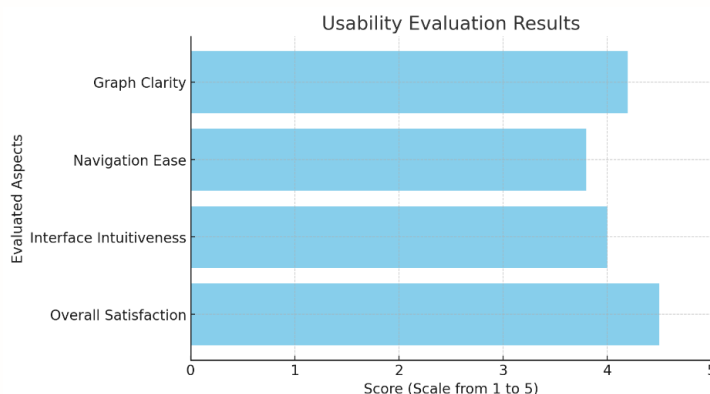


Fig. 9. Usability evaluation results. (Source: Own elaboration).

6.3 Qualitative Analysis

During the interviews, participants highlighted several positive aspects of the system, including the clarity of the graphics and the ease of exporting data. However, they also identified areas for improvement, such as the need for more customization options in visualizations and better organization of interface elements.

Based on these findings, potential improvements for future versions of the system were identified, such as the inclusion of visualization templates and the optimization of control layout.

7 Conclusions

The evaluation demonstrated that the system was well received in terms of graphical clarity and ease of use. However, opportunities for improvement were identified, particularly in customization options and interface element arrangement. These findings will be considered for future iterations of the system design.

7.1 Discussion

The developed interface stood out for its intuitive design, facilitating use across both mobile and desktop devices. The visualization tools enabled users to interpret emotional patterns clearly, making them especially useful for activists.

Additionally, since the system does not store user data, it enhances privacy protection while providing the capability to export results in PDF and CSV formats. This positions the interface as an ethical and practical tool.

However, areas for improvement were identified, such as the incorporation of interactive tutorials for users without prior experience. These features establish the interface as a valuable contribution for analyzing and communicating emotions related to climate change.

7.2 Conclusion

The development of this interface demonstrated how a user-centered approach can lead to technological tools that are both useful and easy to use. Through user feedback, the system was able to adapt to people's needs,

establishing a foundation for future improvements. This process not only highlights the importance of adaptability in design but also how collaboration with end users can enrich the final product.

In future iterations, efforts will focus on expanding testing to more diverse users to ensure accessibility and robustness across different cultural, linguistic, and technical contexts. Additionally, the integration of advanced functionalities will be explored, such as real-time analysis, multilingual support, and personalized result customization, allowing the system to address more complex use cases.

Moreover, the incorporation of interactive data visualization and AI-based predictive methods will be considered, expanding the tool's potential for deeper analysis. Finally, the sustainability and scalability of the system will be evaluated for broader implementation in academic and business environments, further promoting its impact on informed decision-making.

Agradecimientos. La autora principal agradece a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo otorgado para el desarrollo de esta investigación. Asimismo, se reconoce el respaldo del Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX) de la Universidad Autónoma de Zacatecas (UAZ), cuyo acompañamiento académico y recursos tecnológicos hicieron posible la realización de este trabajo.

References

1. Brown, T., Miller, R.: The evolution of sentiment analysis. *Computational Linguistics Review*, Vol. 10, No. 2, pp. 45–67 (2018).
2. Gutiérrez Cervantes, J.C., Lozano Rodríguez, A., Mercado Varela, M.A.: Competencias emocionales en universitarios: un estudio de teoría fundamentada. *Inventio, la génesis de la cultura universitaria en Morelos*, Vol. 20, No. 1, pp. 27–42 (2024).
3. Organizadores de CLIMA NLP.: CLIMATE NLP 2025: The 2nd workshop of natural language processing meets climate change. *Actas del Congreso*, (2025).
4. Doe, J., Smith, A.: Understanding environmental posts: Sentiment and emotion analysis of social media data. *Journal of Environmental Studies*, Vol. 15, No. 3, pp. 123–134 (2020).
5. Hernández Mora, J.J., Sánchez Hernández, M.J., Nava Bautista, H., Collazos Ordóñez, C.A., Hurtado Alegría, J.A., García Toribio, G., Polvo Saldaña, Y.: Measuring the usability of the design of user interface with heuristic evaluation method: two case studies. *Revista Colombiana de Computación*, Vol. 20, No. 1, pp. 23–40 (2019).
6. Guerra, C.A.: Obtención de requerimientos. *Técnicas y estrategia*. SG Software Gu-rú, No. 17 (2006).
7. Hamui-Sutton, A., Varela-Ruiz, M.: La técnica de grupos focales. *Investigación en Educación Médica*, Vol. 2, No. 5, pp. 55–60 (2013).
8. International Organization for Standardization.: ISO/IEC TR 29110-1:2016 – Ingeniería de software y sistemas – Perfiles de ciclo de vida para entidades muy pequeñas (VSEs). ISO, Ginebra (2016).
9. Lasso, J.: Diferentes indicadores para realizar el análisis de sentimiento. Documento en línea. Disponible en: https://phiveres.iniciativas.org/analisis_sentimientos (2024). Accedido el 12 de marzo de 2025.
10. Reyes Podadera, R.: Análisis de sentimiento en redes sociales. Documento en línea. Accedido el 2 de diciembre de 2024 (2023).
11. Quezada, W.: Análisis de sentimientos. Documento en línea. Disponible en: <https://www.wstutdoc.com/ec/document/escuela-superior-politécnica-de-chimborazo/analisis-matematico/analisis-de-sentimientos> (2019).
12. Segado-Bojo, F., Díaz-Campo, J., Navarro-Sierra, N.: Emociones y difusión de noticias sobre el cambio climático en redes sociales. Influencia de hábitos, actitudes previas y usos y gratificaciones en universitarios. *Revista Latina de Comunicación Social*, No. 75, pp. 245–269 (2020).
13. The Bridge.: Diseño emocional: cómo crear interfaces que generen emociones positivas en los usuarios. Documento en línea. Accedido el 12 de marzo de 2025 (2024).
14. Ware, C.: *Information Visualization: Perception for Design*. 2nd edn. Morgan Kaufmann, San Francisco, CA (2004).
15. Zulueta Véliz, Y., Hernández González, A., Despaigne Herrera, E.: Retos en la gestión de los riesgos de proyectos de software. *Revista Cubana de Ciencias Informáticas*, Vol. 2, No. 3–4 (2011).

Aplicación Móvil con Realidad Aumentada para el Centro Histórico de Zacatecas

Karla M. García-Castillo¹, Huizilopoztli Luna-García¹, Roberto Solís-Robles¹,
Hamurabi Gamboa-Rosales¹, Cristian Barria-Huidobro², Klinge O. Villalba-Condori³

¹ Laboratorio de Tecnologías Interactivas y Experiencia del Usuario (LITUX), Universidad Autónoma de Zacatecas,
Jardín Juárez 147, Centro, Zacatecas 98000, México.

² Centro de Investigación en Ciberseguridad, Universidad Mayor de Chile, Manuel Montt 367, Providencia 7500628,
Chile.

³ Vicerrectorado de Investigación, Universidad Católica de Santa María, Yanahuara 04013, Perú.

karla.garcia.castillo, hlugar, rsolis, hamurabigr@uaz.edu.mx

cristian.barria@umayor.cl

kvillalba@ucsm.edu.pe

Resumen. Este artículo presenta el desarrollo de un prototipo de aplicación móvil de Realidad Aumentada (RA) para la visualización de información sobre edificios históricos en el centro de la ciudad de Zacatecas, siguiendo un enfoque de Diseño Centrado en el Usuario (DCU). El objetivo fue diseñar una herramienta interactiva que enriqueciera la experiencia cultural del usuario al proporcionar datos sobre monumentos y edificaciones, además de una agenda actualizada de eventos culturales cercanos. Se definió el perfil del usuario potencial mediante entrevistas y observaciones, lo que permitió adaptar el diseño del prototipo a sus necesidades y expectativas. Una vez desarrollado, se realizaron pruebas con usuarios reales, implementando el Sistema de Escalas de Usabilidad (SUS) para evaluar el prototipo, obteniendo una calificación de 71.5 puntos, posicionando el producto como aceptable. Además, se recopiló retroalimentación cualitativa por parte de los usuarios que permitió identificar mejoras clave para la evolución del prototipo hacia una aplicación funcional. Pese a que los resultados son promisorios, se reconocen las limitaciones técnicas asociadas a la implementación de RA y el alcance de las pruebas realizadas, lo que refuerza la necesidad de futuras iteraciones. El prototipo demuestra potencial para fortalecer el turismo cultural y fomentar la valoración del patrimonio local.

Keywords: Prototipo de Aplicación Móvil, Realidad Aumentada, Historia, Turismo, Diseño Centrado en el Usuario.

1 Introducción

Los avances tecnológicos han transformado múltiples actividades humanas, entre ellas el turismo, el cual ha integrado progresivamente las Tecnologías de la Información y la Comunicación (TIC) para mejorar la experiencia del turista, desde la planeación de su viaje hasta la interacción con el entorno que visita. Entre las tecnologías emergentes, la Realidad Aumentada (RA) destaca por su capacidad de presentar información digital en tiempo real sobre el entorno físico del usuario, enriqueciendo su experiencia sin reemplazar la realidad que percibe. Este tipo de tecnología combina cámaras, sensores y pantallas de visualización para superponer los elementos digitales, se implementa frecuentemente en aplicaciones móviles como mapas con indicaciones visuales, videojuegos o herramientas de diseño de interiores. Su funcionamiento puede clasificarse en dos tipos principales: RA basada en marcadores, que utiliza códigos QR o imágenes específicas para activar contenido digital, y RA sin marcadores, que emplea sensores como GPS o acelerómetros para mapear el entorno y ubicar el contenido digital de manera automática [1].

El centro histórico de Zacatecas se caracteriza por su riqueza arquitectónica, conformada por plazas, jardines y monumentos de diversos estilos y épocas. Esta zona representa un importante atractivo turístico y componente económico para el estado. De acuerdo al Tercer informe de Gobierno (2024), la zona conurbada Guadalupe-Zacatecas registró una ocupación hotelera de 624,928 personas, de las cuales 113,582 tenían como objetivo de visita museos y edificios históricos. Asimismo, se realizaron nueve festivales y más de 400 eventos artístico-

culturales en el centro histórico con una audiencia de 416,300 asistentes [2]. Estas cifras ilustran la relevancia de la zona como eje cultural y turístico.

Este trabajo propone el desarrollo de un prototipo de Aplicación Móvil que integra RA basada en marcadores, diseñada bajo un enfoque del Diseño Centrado en el Usuario (DCU). La elección de esta tecnología responde a criterios de simplicidad y control durante las fases iniciales de diseño y evaluación, no obstante, se contempla la incorporación de RA sin marcadores en versiones futuras, lo que permitiría posicionar automáticamente el contenido digital sin depender de elementos visuales específicos.

El objetivo del prototipo es ofrecer una herramienta interactiva que permita visualizar información relevante sobre edificios históricos en el centro de Zacatecas, así como consultar una agenda de eventos culturales cercanos, enriqueciendo la experiencia del visitante y fomentando el interés del público local por su patrimonio cultural.

2 Estado del arte

La RA ha sido ampliamente adoptada en aplicaciones móviles, especialmente en redes sociales a través de filtros interactivos en plataformas como Instagram, Snapchat, TikTok y Facebook. Éstos presentan una gran variedad de elementos y efectos digitales que se superponen en tiempo real a la visión captada por la cámara del dispositivo móvil, sus funciones van desde efectos cosméticos hasta proyección de objetos, animaciones y entornos virtuales. [3].

En el rubro del entretenimiento, diversas compañías han incorporado RA para ofrecer nuevas formas de interacción con sus productos. Un caso ampliamente reconocido es el videojuego Pokémon GO, desarrollado por Niantic en 2016, el cual introdujo un modo de RA+ en el que criaturas y objetos virtuales aparecen ancladas al entorno real, utilizando la ubicación geográfica del jugador [4]. Más allá del componente lúdico, estudios han documentado su impacto en la conducta de los usuarios. Por ejemplo, Lawler-Sagarin (2016), en su trabajo *Enhanced Community Through Augmented Reality: Social Benefits of Pokémon Go*, reportó que el 78% de los usuarios exploró nuevos lugares gracias a la aplicación y el 35% descubrió nuevos negocios locales, evidenciando el potencial de esta tecnología para fomentar la exploración del entorno físico y fortalecer el vínculo con la comunidad local [5].

En el sector comercial, aplicaciones como IKEA Place [6] y Room Decorator de Amazon [7] permiten al usuario visualizar modelos tridimensionales a escala real de productos del catálogo en su espacio físico, para evaluar la disposición, tamaño y estética del objeto seleccionado, apoyando la toma de decisiones de compra.

En el ámbito académico, la Universidad Nacional de San Juan, Argentina, desarrolló una aplicación educativa, que a través de tecnología RA permite visualizar iconografía tridimensional con información técnica y contextual de sitios con vulnerabilidad sísmica en un edificio de uso público, con el propósito de promover la concientización de riesgos estructurales entre los usuarios [8].

Estos ejemplos muestran la versatilidad de la RA y su capacidad de aplicarse a diferentes contextos. El presente trabajo propone una aplicación que combina componentes informativos, culturales y de asistencia, integrando el enfoque del DCU para construir una experiencia orientada a satisfacer las necesidades del usuario final, mejorando la experiencia de visitantes y habitantes locales en el centro histórico de Zacatecas.

3 Metodología: Diseño centrado en el usuario

El desarrollo de este proyecto se llevó a cabo siguiendo el enfoque del Diseño Centrado en el Usuario (DCU), el cual prioriza identificar y atender las características, necesidades y expectativas de éste a lo largo de todo el proceso de diseño. Éste enfoque promueve la participación activa de los usuarios desde las etapas iniciales, permitiendo realizar ajustes iterativos basados en sus retroalimentaciones. Aunque no existe una guía definitiva que dicte el proceso completo del DCU, es posible apoyarse en lineamientos y normas internacionales. En este

caso, se adoptan los principios establecidos en el estándar ISO 9241-210:2010, el cual establece un marco para diseñar sistemas interactivos usables y accesibles [9].

El Diseño Centrado en el Usuario se puede dividir en cuatro etapas fundamentales:

1. Especificación del contexto de uso.
2. Especificación de los requisitos.
3. Producción de soluciones de diseño.
4. Evaluación del diseño.

Éstas se aplican de forma cíclica permitiendo iteraciones continuas para afinar el prototipo conforme se obtiene nueva información [Fig.1] [10].

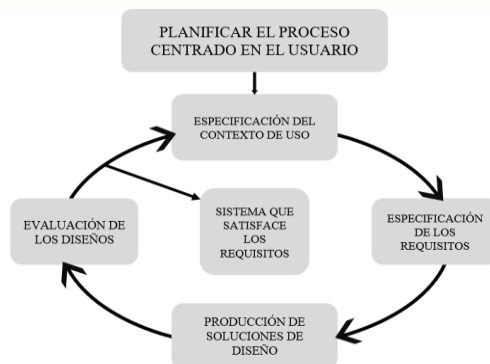


Fig. 1. Etapas del diseño centrado en el usuario [10].

A continuación, se describen las actividades realizadas en cada una de las etapas mencionadas, con base en la aplicación práctica de la metodología.

3.1 Especificación del contexto de uso

En esta etapa se busca identificar y analizar los usuarios potenciales, así como de las condiciones en las que interactuarían con el sistema. El objetivo fue comprender el perfil, necesidades, intereses y hábitos de los usuarios prospecto. Para ello, se diseñaron y aplicaron encuestas en línea a través de la plataforma Google Forms. Se elaboraron dos formatos para contemplar las posibilidades de entrevistar tanto a individuos locales como foráneos, ambos formularios incluyeron preguntas orientadas a recopilar información de las siguientes categorías: aspectos demográficos, motivaciones y objetivos de la visita, comportamiento y preferencias de uso, tecnología y usabilidad, preferencias de contenido y expectativas de la aplicación.

El cuestionario de contexto local incluye preguntas adicionales acerca de experiencias cotidianas en el centro histórico [11] [12]. Con la información recopilada, se implementó la “técnica persona”, método propuesto por Alan Cooper (1983) que consiste en crear perfiles ficticios representativos basados en el análisis cualitativo de usuarios reales. Estos perfiles permiten desarrollar un diseño orientado hacia las necesidades y expectativas de “un usuario tipo” específico [13].

3.2 Especificación de los requisitos

Una vez definido el contexto de uso y comprendido el perfil de los usuarios objetivo, se procedió a identificar los requerimientos funcionales y no funcionales del prototipo, éstos se refieren a las capacidades esenciales que el prototipo debe ofrecer para que el usuario pueda cumplir sus objetivos. Para determinarlos, se tomaron como referencia los patrones detectados en las encuestas, elaborando gráficos estadísticos para facilitar su interpretación.

3.3 Producción de soluciones de diseño

Con los requisitos definidos, se procedió a crear propuestas de diseño orientadas a materializar las funcionalidades del sistema. En las primeras fases del desarrollo se implementó la técnica de Paper Prototyping, metodología de baja fidelidad que permite representar visualmente la estructura, navegación y funcionalidad propuesta de la aplicación sin necesidad de recursos técnicos complejos [14].

Posteriormente, se desarrolló un prototipo de alta fidelidad utilizando la herramienta Justinmind [15], seleccionada por su capacidad para simular de forma dinámica las interacciones del sistema, incluyendo transiciones de pantallas, respuestas táctiles y flujos de navegación realistas. Esta etapa es resultado de la primera iteración del ciclo continuo de mejora propuesto por el DCU.

3.4 Evaluación del diseño

Una vez desarrollado el prototipo de alta fidelidad, se evaluó con usuarios reales para medir su usabilidad y detectar áreas a mejorar. Se aplicó el Sistema de Escalas de Usabilidad, SUS (por sus siglas en inglés), una herramienta estandarizada que permite obtener una puntuación cuantitativa sobre la percepción del sistema en términos de eficacia, eficiencia y satisfacción [16]. Además, se recopiló retroalimentación individual de los usuarios sobre su experiencia visual y funcional con el prototipo.

4 Resultados

La implementación del DCU, permitió identificar al público objetivo, conocer necesidades y expectativas, definir los requerimientos del prototipo y desarrollar un producto. A continuación, se presentan los resultados obtenidos en cada etapa del proceso.

4.1 Definición del contexto de uso

La encuesta aplicada a través de Google Forms contó con la participación de 63 personas locales y 20 foráneas. El análisis de los datos permitió identificar patrones clave para delimitar al usuario objetivo, edad, ocupación, motivaciones y familiaridad con aplicaciones móviles.

Los resultados mostraron que el grupo más representativo corresponde a personas entre 25 y 34 años, en su mayoría con ocupación laboral activa y conocimientos intermedios o avanzados en el uso de dispositivos móviles y aplicaciones de tecnología similar a la propuesta en el prototipo. En cuanto al estilo de visita, tanto locales como foráneos manifiestan preferencia por recorridos a pie, valoran el aprendizaje cultural y muestran interés por herramientas que les permitan optimizar el tiempo durante sus visitas [17].

La elección de este segmento etario como público objetivo se justifica por su alta representatividad en la muestra al ser el sector con mayor disposición a participar y demostrar un interés genuino en las propuestas tecnológicas vinculadas al patrimonio cultural. Su familiaridad con interfaces digitales y apertura hacia nuevas experiencias, permiten prever una interacción eficaz con el prototipo sin requerir una curva de aprendizaje pronunciada.

4.2 Perfil de usuario

Con los patrones detectados, se construyó un perfil persona que representa las características y contexto del público objetivo. Este perfil describe a un usuario interesado en explorar nuevos lugares, con apreciación por la arquitectura y cultura local, que busca soluciones tecnológicas que se integren de forma sencilla y efectiva a su experiencia turística [18].

El tiempo libre es valioso y limitado debido sus responsabilidades laborales, lo cual refuerza la necesidad de diseñar un producto eficiente, intuitivo y de fácil acceso. Estos elementos fueron clave para orientar decisiones de diseño y funcionalidad del prototipo.

4.3 Especificación de requerimientos

El análisis de los apartados correspondientes a tecnología y usabilidad, contenido y expectativas, permitió delimitar las características técnicas y funcionales que debía cubrir el prototipo.

En primer lugar, se identificó preferencia hacia el sistema operativo Android, así como el uso de aplicaciones de navegación como Google Maps [Fig.2]. Esta información orientó el desarrollo inicial hacia diseños de entornos con los que el usuario esté familiarizado.

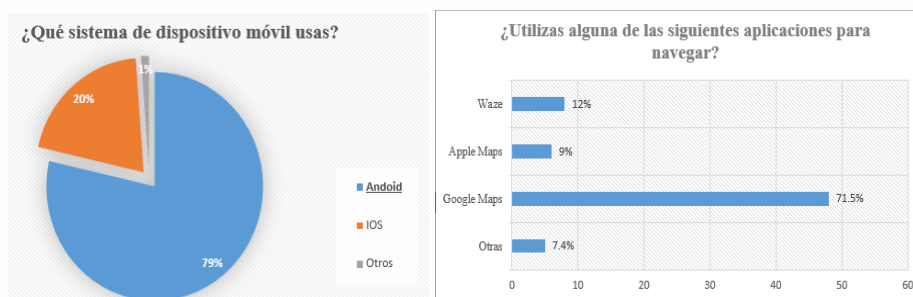


Fig. 2. Tecnología y usabilidad, preferencias de sistema móvil y navegación.

Por otra parte, una porción significativa de los participantes manifiesta haber usado previamente aplicaciones con herramientas de RA, principalmente a través de filtros interactivos en redes sociales [Fig. 3]. Esto sugiere un panorama favorable para la adopción de esta tecnología en el prototipo.

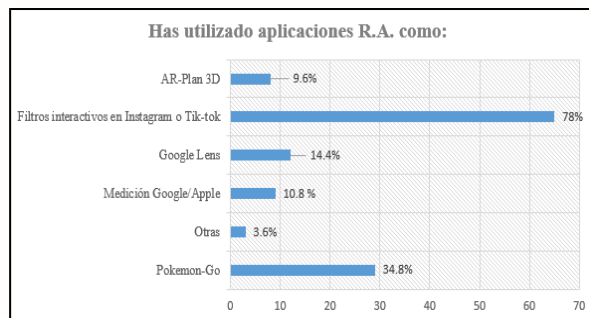


Fig. 3. Tecnología y usabilidad, resultados sobre experiencias previas con RA.

Respecto al tipo de información que los usuarios desean recibir, los resultados indicaron mayor interés por datos prácticos como: información básica sobre el edificio o monumento, sus horarios de apertura y cierre, así como acontecimientos históricos relevantes [Fig. 4]. Esta preferencia fundamentó el incluir fichas informativas dentro de la experiencia RA.

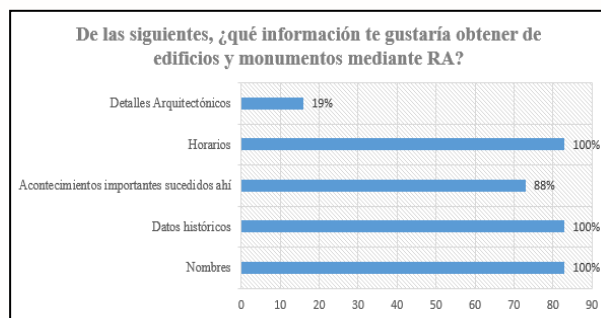


Fig. 4. Preferencias de contenido RA.

En cuanto a la experiencia de uso, se obtuvieron respuestas positivas hacia el incentivo mediante premios o recompensas y se mostró interés en funciones adicionales como notificación de eventos locales [Fig.5].

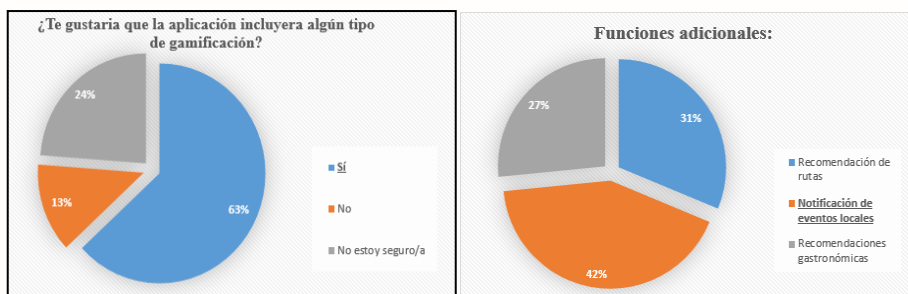


Fig. 5. Expectativas del prototipo y funciones adicionales.

Con la información recolectada, se elaboró un documento de especificación de requisitos para el prototipo [19], siguiendo un formato respaldado en el estándar IEEE-830 [20]. En él se definieron tanto los requerimientos funcionales, como la visualización de en RA, integración de mapas y alertas, además de los no funcionales, como compatibilidad, eficiencia y facilidad de uso.

4.4 Producción de soluciones de diseño

Con los requerimientos definidos se procedió a la elaboración de propuestas de diseño, enfocadas en representar de manera visual la funcionalidad y estructura del sistema. Esta etapa fue ejecutada bajo el principio iterativo del DCU, permitiendo validar ideas tempranas y realizar ajustes progresivos con iteraciones subsecuentes.

Durante la primera iteración, se implementó la técnica Paper Prototyping, en la cual se presentaron las propuestas de una pantalla principal, ubicación de logotipos, ventana de vista mapa, pantalla de escaneo mediante cámara para buscar elementos RA y la disposición de fichas informativas. Este prototipo fue acompañado de un esquema ilustrativo con la descripción de las funciones propuestas, permitiendo identificar las áreas a mejorar [Fig. 6].

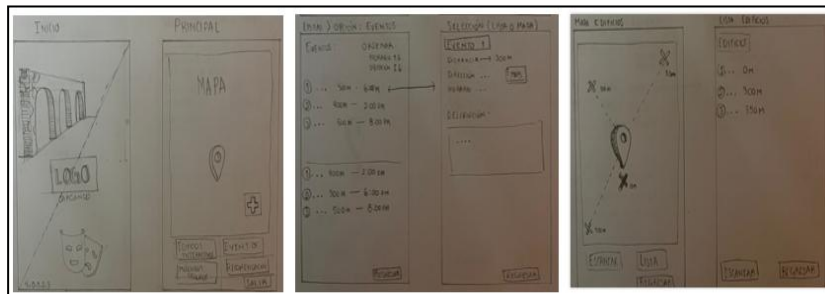
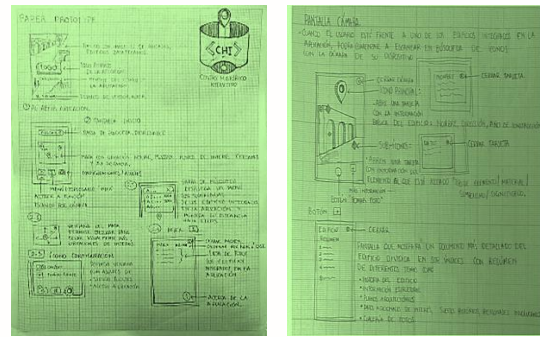


Fig. 6. Paper prototype.

Posteriormente, se desarrolló un prototipo de alta fidelidad con la herramienta Justinmind [Fig.7]. Esta versión integró elementos visuales más precisos además de una simulación de navegación visual y táctil. El diseño se elaboró conservando los principios de:

- Simplicidad: priorizando una navegación clara y accesible desde la pantalla principal.
- Intuición: utilizando íconos y etiquetas familiares para los usuarios, inspirados en aplicaciones de navegación más utilizadas.
- Accesibilidad al contenido: propuesta de información clave sobre los edificios, presentada de forma concisa evitando la saturación visual.

Gracias al enfoque iterativo del DCU, la transición del prototipo en papel al de alta fidelidad permitió validar elementos visuales y funcionales. Esta etapa resultó fundamental para consolidar la propuesta alineada a las necesidades del público objetivo.

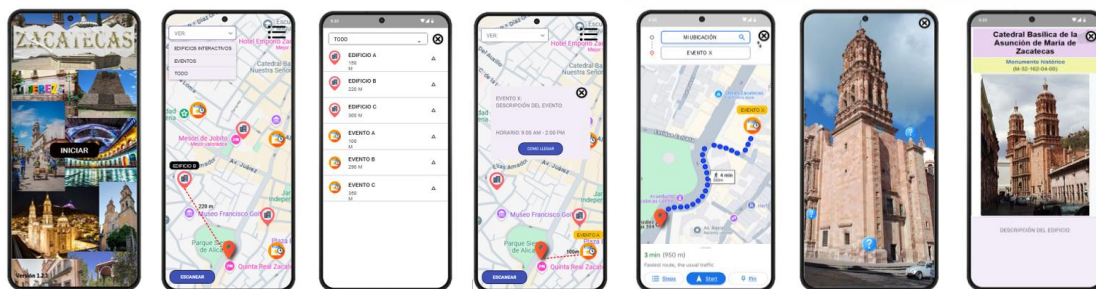


Fig. 7. Prototipos de alta fidelidad.

4.5 Evaluación de los diseños

Siguiendo el enfoque iterativo del DCU y la teoría de saturación de usabilidad de Nielsen [21], se seleccionaron 5 usuarios con características acorde al perfil persona para la primera prueba del prototipo. Tras la experiencia, se aplicó la evaluación SUS y se obtuvo una calificación promedio de 71.5, que en escala del sistema, clasifica el prototipo como aceptable dentro de la escala de usabilidad.

Tabla 1. Resultados de la evaluación de usabilidad (SUS).

Usuario	Puntaje SUS (0-100)	Calificación interpretativa
U1	75	Aceptable
U2	82.5	Aceptable
U3	67.5	Marginal
U4	72.5	Aceptable
U5	60	Marginal
Promedio:	71.5	Aceptable

Los resultados individuales de las encuestas se recopilaron en un documento disponible como Anexo 1 [22], por motivos de privacidad, se ha cubierto el nombre de los participantes.

La retroalimentación de los participantes permitió identificar áreas de mejora:

Visuales: Estilización de los íconos, uso de fuentes de texto más atractivas y aumento de transparencia en los menús desplegables.

Funcionales: Mostrar el botón de escanear solo frente a edificios interactivos, mejorar la organización de la información en listas, y permitir la interacción tanto en el mapa como en la vista de cámara.

5 Discusión

En este trabajo, la aplicación del enfoque de Diseño Centrado en el Usuario (DCU) brindó una guía jerárquica lógica y efectiva que permitió identificar al público objetivo, entender sus necesidades, y adaptar la propuesta de diseño en función de sus hábitos, intereses y contexto de uso. Este enfoque fue clave para diseñar desde una perspectiva empática y funcional, reflejándose en la buena recepción del prototipo y su calificación aceptable en usabilidad desde la primera interacción con un público real.

Los resultados coinciden con estudios previos que destacan el potencial de las aplicaciones con funciones de RA para fomentar la exploración del entorno físico, como es el caso de Pokémon GO, donde los usuarios reportaron mayor disposición a recorrer espacios urbanos. En este proyecto los participantes también mostraron una actitud positiva hacia el uso de RA en un contexto cultural, motivación que, de ser impulsada adecuadamente, podría ofrecer resultados similares a los de otras aplicaciones.

El proceso iterativo permitió detectar áreas de mejora gracias a la retroalimentación de los usuarios, como ajustes visuales, mejoras en la navegación y optimización del contenido. Estos comentarios son valiosos para futuras versiones y demuestran el valor de involucrar al usuario desde etapas tempranas.

Aunque el prototipo se diseñó principalmente para un público joven-adulto, se identifica una oportunidad de adaptarlo al ámbito educativo para atender a un público estudiantil, teniendo fines didácticos para promover el conocimiento del legado histórico de su ciudad a edad temprana.

Es importante reconocer las limitaciones del estudio, ya que el prototipo aún no cuenta con funcionalidad completa y las pruebas se realizaron en un entorno simulado, por lo que se requerirán evaluaciones en contexto real conforme avancen las iteraciones del producto. Pese a las limitaciones, el estudio sienta una base sólida para continuar hacia el desarrollo de una aplicación funcional. El interés genuino mostrado por los usuarios,

sumado al creciente uso de la RA en diversas áreas, sugiere un camino prometedor para el diseño de herramientas digitales que contribuyan al enriquecimiento cultural, la accesibilidad al conocimiento y la mejora de experiencias turísticas.

6 Conclusiones

La aplicación del DCU resultó ser una excelente estrategia para el desarrollo del prototipo. Gracias a éste logramos un puntaje aceptable en usabilidad desde la primera iteración. La recepción inicial de la propuesta fue positiva, indicando que la información cultural proporcionada a través de RA tiene un gran potencial. La retroalimentación de los usuarios señaló áreas claves para la mejora iterativa derivada del prototipo, lo que permitirá como trabajo futuro avanzar hacia una aplicación final completamente funcional.

Agradecimientos. La autora principal agradece a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo otorgado para el desarrollo de esta investigación. Asimismo, se reconoce el respaldo del Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX) de la Universidad Autónoma de Zacatecas (UAZ), cuyo acompañamiento académico y recursos tecnológicos hicieron posible la realización de este trabajo.

Referencias

1. IBM. (2024). *Augmented Reality*. IBM. <https://www.ibm.com/mx-es/topics/augmented-reality>.
2. Gobierno del Estado de Zacatecas: *3er Informe de Gobierno 2021-2027*. Gobierno del Estado de Zacatecas, Zacatecas (2024). Sección 2.6 Política Pública, pp.199-205. <https://3erinformedavidmonreal.zacatecas.gob.mx/archivos/3ERINFORME2024.pdf>, last accessed 2024/12/12.
3. Virtualmedia XR: *¿Qué son los Filtros de Realidad Aumentada?*, <https://www.virtualmediaxr.com/filtro-realidad-aumentada/>, last accessed 2024/12/12.
4. Dorward, L.J., Mittermeier, J.C., Sandbrook, C., Spooner, F. (2017). Pokémon Go: Benefits, Costs, and Lessons for the Conservation Movement. *Conservation Letters*, 10(3), 160–165. <https://doi.org/10.1111/conl.12326>.
5. Lawler-Sagarin, K.A., Sagarin, B.J., & Pederson, A. (2023). Enhanced Community Through Augmented Reality: Social Benefits of Pokémon Go. *Psychological Reports*. <https://doi.org/10.1177/00332941231197155>.
6. IKEA. (2017). *IKEA Place app launched to help people virtually place furniture at home*. <https://www.ikea.com/global/en/newsroom/innovation/ikea-launches-ikea-place-a-new-app-that-allows-people-to-virtually-place-furniture-in-their-home-170912>, last accessed 2024/12/12.
7. Perez, S. (2020). *Amazon rolls out a new AR shopping feature for viewing multiple items at once*. TechCrunch. <https://techcrunch.com/2020/08/25/amazon-rolls-out-a-new-ar-shopping-feature-for-viewing-multiple-items-at-once/>, last accessed 2025/04/17.
8. Heredia, S.F. (2024). *Tecnologías digitales. Desarrollo de aplicación de realidad aumentada como herramienta de visualización de vulnerabilidad sísmica edilicia*. Tecnología y Ciencia. <https://dx.doi.org/10.33414/rtyc.50.1-17.2024>
9. Online Browsing Platform (OBP). (2010). *ISO 9241-210:2010 Ergonomics of human-system interaction—Part 210: Human-centered design for interactive systems*. International Organization for Standardization. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en>.
10. Maguire, M. C. (1995). *RESPECT User-Centered Requirements Handbook* (p. 20).
11. Castillo, K. M. (2024). *Encuesta para prototipo de recorridos turísticos con RA (foráneos)* Google Forms. <https://forms.gle/JCYksbUqP9XMPWWXA>.
12. Castillo, K. M. (2024) *Encuesta para prototipo, recorriendo el centro histórico con RA, (locales)* Google Forms <https://forms.gle/NRVb4ZR2kgxkcpcdk8>.
13. Design Thinking. *¿Qué es la técnica “personas”?* DesignThinking.es. <https://designthinking.es/personas/>.
14. Sefelin, R., Dörflinger, T., & Tscheligi, M. (2003). *Paper prototyping: What is it good for? A comparison of paper- and computer-based low-fidelity prototyping*. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* pp. 778–779. ACM. <https://doi.org/10.1145/765891.765986>.
15. Justinmind. (2024). *Justinmind*. <https://justinmind.com/>, last accessed 2024/12/12.
16. Brooke, J. (1996). SUS: A quick and dirty usability scale. https://www.researchgate.net/publication/228593520SUS_A_quick_and_dirty_usability_scale.
17. Castillo, K. M. (2025). *Resultados de encuestas, Google Drive*. https://drive.google.com/file/d/1mViYmsDlmd_38fKo0wPQ78vUv-gGZm9W/view?usp=sharing.

18. Castillo, K. M. (2025). *Perfil persona*. Google Drive.
<https://drive.google.com/file/d/1nzPwXGKvx95t2ioy0Pep0eQdM4dPiOOh/view>.
19. Castillo, K. M. (2025). *Documento de requerimientos*. Google Drive.
<https://drive.google.com/file/d/1XQ1ls28tQppjJR4Qj1qjLkqcYjhlvzh8/view?usp=sharing>.
20. IEEE. (2005). *IEEE 830-1998 - IEEE Recommended Practice for Software Requirements Specifications*.
<https://www.fdi.ucm.es/profesor/glmenendez/docs/is0809/ieee830.pdf>
21. Nielsen, J. (2000). Why you only need to test with 5 users. Nielsen Norman Group.
<https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>.
22. Castillo, K. M. (2025). Anexo 1: resultados de evaluaciones. Google Drive.
https://drive.google.com/file/d/1LahOiliE_m26WIAz01i9ZzF1dl3cjNQ-/view?usp=sharing.

Diseño de una Aplicación Educativa en Realidad Virtual para Fomentar Habilidades Críticas en el Pensamiento Computacional

Isaías R. Valdivia¹, Huizilopoztli Luna-García¹, José M. Celaya-Padilla¹, Wilson J. Sarmiento², Carlos H. Espino-Salinas¹, Klinge O. Villalba-Condori³

¹ Laboratorio de Tecnologías Interactivas y Experiencia de Usuario, Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Zacatecas 98000, México

² Ingeniería en Multimedia, Universidad Militar de Nueva Granada, Cra 11, Bogotá 101-80, Colombia

³ Vicerrectorado de Investigación, Universidad Católica de Santa María, Arequipa 04002, Perú
isaia.valdivia@uaz.edu.mx, hlugar@uaz.edu.mx, jose.celaya@uaz.edu.mx, wilson.sarmiento@unimilitar.edu.co, carlospino@uaz.edu.mx, kvillalba@ucsm.edu.pe

Resumen. El pensamiento computacional es una competencia clave en la educación del siglo XXI, esencial para la resolución de problemas de manera estructurada en diversas disciplinas. Sin embargo, su enseñanza se presenta tradicionalmente con un enfoque teórico o con actividades que no aprovechan todo el potencial de las nuevas tecnologías en la educación. La realidad virtual ha demostrado ser muy eficaz para mejorar el aprendizaje, al proporcionar entornos inmersivos e interactivos, en los que los estudiantes pueden desarrollar sus habilidades mediante la participación activa. Este artículo presenta el diseño de Revipec, una aplicación de realidad virtual orientada al fortalecimiento de las habilidades del pensamiento computacional en estudiantes de secundaria. La aplicación sigue la metodología MEDEERV, la cual permite estructurar el diseño en tres fases: diseño sistemático de la instrucción, diseño funcional e implementación. A través de esta metodología, se asegura una correcta alineación pedagógica con las experiencias interactivas de la realidad virtual.

Palabras Clave: Pensamiento Computacional, Realidad Virtual, Educación Secundaria, Aprendizaje Inmersivo.

1 Introducción

El Pensamiento Computacional (PC) es una competencia clave en la educación del siglo XXI, la cual es comparable con otras como la lectura, escritura y matemáticas [1]. El enfoque seguido en el PC permite a los estudiantes resolver problemas de una manera estructurada a través de la descomposición, abstracción, pensamiento algorítmico y evaluación de soluciones [2]. Sin embargo, tradicionalmente su enseñanza se basa en métodos teóricos y actividades digitales convencionales, lo que puede llegar a reducir la motivación y el compromiso de los estudiantes [3].

Las tecnologías inmersivas, como la Realidad Virtual (RV), han demostrado ser herramientas capaces de transformar la educación al crear entornos interactivos y experienciales [4]. La implementación de RV en la educación ha demostrado potenciar la motivación y facilitar el aprendizaje activo [5].

Este artículo presenta el diseño de Revipec, una Aplicación de RV que fortalece el PC en estudiantes de secundaria. Este diseño se basa en la Metodología para el Desarrollo de Sistemas Educativos de Realidad Virtual (MEDEERV) [6], asegurando una correcta alineación pedagógica con experiencias interactivas.

2 Fundamentación teórica

El término Pensamiento Computacional (PC) fue popularizado por Jeanette Wing en 2006, definiéndolo como el proceso de formular problemas de tal manera que una computadora pueda resolverlos eficientemente [7]. Tradicionalmente, la adquisición de estas habilidades ha dependido principalmente de enfoques teóricos o

herramientas digitales convencionales, lo que puede impactar negativamente en la motivación de los estudiantes.

La Realidad Virtual (RV) se ha consolidado como una herramienta muy eficaz en entornos educativos al proporcionar entornos inmersivos que facilitan la exploración y experimentación de conceptos y habilidades complejas [8]. En el contexto del PC, la RV permite a los estudiantes manipular elementos virtuales de manera intuitiva, favoreciendo el aprendizaje y la retención de conocimientos [9].

El uso de la RV en específico para enseñar el PC, ha sido explorado en otros trabajos relevantes. Por ejemplo, Agbo et al. [10] desarrollaron iThinkSmart, una aplicación educativa en RV basada en minijuegos, orientada a conceptos como recursividad, búsqueda de soluciones óptimas y pensamiento algorítmico. Los resultados mostraron mejoras significativas en competencias del PC, además de una mayor motivación de los estudiantes frente a métodos tradicionales.

Por su parte, Zhao y Shute [11], evaluaron Penguin Go, un videojuego educativo que utiliza bloques de programación visual y una narrativa interactiva para guiar a los estudiantes en la resolución de retos. Aplicado en estudiantes de secundaria, su uso mostró mejoras significativas en habilidades como el pensamiento algorítmico, la lógica condicional y la optimización de soluciones.

Asimismo, Zapata-Cáceres et al. [12] presentaron Blue Ant Code, una herramienta cuyo enfoque se basa en la programación visual con misiones secuenciales, diseñada para desarrollar las habilidades del PC mediante desafíos progresivos. Su estudio reveló mejoras significativas en el rendimiento de los estudiantes frente a métodos tradicionales, así como una retención del aprendizaje comprobada semanas después de la intervención.

Estos estudios ofrecen un respaldo teórico y empírico al uso de tecnologías inmersivas e interactivas para el fortalecimiento de las habilidades del PC, aportando referentes clave que han influido en el diseño de Revipec. A partir de los resultados obtenidos por ellos, se tomaron decisiones como la estructuración progresiva de misiones, el uso de una narrativa inmersiva y la integración de elementos de gamificación.

En el contexto de este trabajo, se adopta una perspectiva funcional del PC, centrada en cinco habilidades fundamentales: descomposición de problemas, abstracción, pensamiento algorítmico, evaluación de soluciones y reconocimiento de patrones. Estas están ampliamente reconocidas en la literatura [2][13], orientan el diseño instruccional de las actividades y la estructuración de los desafíos en la aplicación. De esta forma, se asegura el desarrollo progresivo y contextualizado de estas habilidades dentro de un entorno inmersivo.

Con base en este marco conceptual y en los hallazgos revisados, se estructuró el diseño de Revipec siguiendo la metodología de MEDEERV, que estructura el desarrollo en tres fases principales: diseño sistemático de la instrucción, diseño funcional e implementación. Esta metodología permite diseñar una experiencia educativa alineada con objetivos pedagógicos específicos, así como integrar las herramientas que los entornos inmersivos brindan para generar una experiencia de aprendizaje efectiva.

3 Metodología de diseño

El desarrollo de Revipec se basa en un enfoque instruccional estructurado, que integra la Realidad Virtual (RV) para fortalecer el PC en estudiantes de secundaria. Se eligió la Metodología para el Desarrollo de Experiencias Educativas en Realidad Virtual (MEDEERV) por su enfoque integral que articula el diseño pedagógico con los aspectos técnicos y funcionales propios de entornos inmersivos. Esta metodología ha sido aplicada con éxito para desarrollos centrados en el aprendizaje mediante dispositivos móviles que hacen uso de la RV, lo que la convierte en una opción que garantiza la coherencia entre los objetivos educativos, mecánicas de interacción y contenido.

Una de las principales fortalezas de MEDEERV es su primera fase, el *Diseño Sistemático de la Instrucción*, donde se realiza un análisis instruccional detallado que parte del diagnóstico del contexto educativo y de las habilidades cognitivas que se busca desarrollar, estableciendo objetivos pedagógicos concretos. Esto asegura que el entorno virtual no sea solo funcional o atractivo visualmente, sino que esté fundamentado en principios instruccionales claros. De esta manera, el entorno responde a un propósito educativo propiamente definido,

permitiendo al estudiante aprender a través de experiencias diseñadas con base en procesos cognitivos específicos.

La segunda etapa, el *Diseño Funcional*, se centra en definir las interacciones, recursos y dinámicas del entorno virtual, mientras que la tercera, la *Implementación*, contempla la construcción del entorno, así como su validez y ajustes según los objetivos de aprendizaje. Esta secuencia asegura que el desarrollo de la experiencia inmersiva mantenga una coherencia pedagógica desde la planificación hasta la ejecución.

3.1 Diseño sistemático de la instrucción

La fase de diseño instruccional partió del reconocimiento de una necesidad educativa actual: el PC se ha identificado como una competencia fundamental en la era digital, siendo equiparable a otras habilidades como la lectura, la escritura y las matemáticas [14]. Sin embargo, su integración formal en los currículos escolares, especialmente en las primeras fases educativas, aún es muy limitada en muchos contextos educativos [15]. Esta brecha entre la demanda social y la oferta formativa actual motivó el enfoque de Revipec como una herramienta complementaria para el desarrollo de estas habilidades mediante experiencias prácticas e inmersivas.

Esta fase se enfocó en la selección de actividades validadas en la enseñanza del Pensamiento Computacional (PC) y su adaptación a un entorno de RV. Para esto se priorizaron las habilidades de descomposición de problemas, abstracción, pensamiento algorítmico, evaluación de soluciones y reconocimiento de patrones.

En lugar de presentar los problemas como actividades aisladas, Revipec utiliza un enfoque narrativo en el que los estudiantes asumen el rol de un investigador que debe resolver una serie de desafíos. La progresión está diseñada para reforzar las habilidades de forma gradual mediante el análisis de patrones, optimización de acciones y resolución de problemas en un entorno interactivo.

3.2 Diseño funcional

Revipec fue desarrollado en Unity 3D, utilizando el visor de RV Meta Quest 2, en conjunto con los Oculus Touch Controllers incluidos con el mismo para permitir interacciones intuitivas. El entorno 3D facilita la exploración y manipulación de elementos virtuales, para tener una experiencia inmersiva.

La gamificación se integra a través de desafíos intuitivos que incentivan el compromiso del estudiante, mientras que la narrativa es el hilo conductor que vincula cada actividad para motivarlo.

3.3 Implementación

Actualmente, Revipec se encuentra en fase de desarrollo, con un enfoque iterativo que busca optimizar la alineación con los objetivos educativos. Se tiene planeado la realización de pruebas piloto en entornos educativos controlados, con el fin de recopilar datos empíricos sobre la efectividad de la aplicación en el desarrollo de las habilidades del PC.

4 Descripción de la aplicación Revipec

Revipec es una aplicación diseñada para fortalecer las habilidades del Pensamiento Computacional (PC) en estudiantes de secundaria. Se basa en la adaptación de desafíos validados en competencias Bebras, con una reestructuración para una resolución interactiva en Realidad Virtual (RV).

4.1 Estructura y narrativa

El juego se compone de siete misiones, cada uno presenta un puzzle a resolver haciendo uso de diferentes habilidades. Los jugadores se enfrentan a problemas progresivos que requieren descomposición de problemas, abstracción, pensamiento algorítmico, evaluación de soluciones y reconocimiento de patrones.



Fig. 1. Ejemplos de niveles creados para la aplicación Revipec, donde el jugador tiene que encontrar el túnel de ingreso correcto (izquierda) y donde el jugador debe identificar patrones de movimiento en los enemigos (derecha).

La elección de siete misiones responde a un equilibrio entre la cobertura de habilidades cognitivas clave y la viabilidad técnica del desarrollo. Cada misión fue diseñada para abordar por lo menos una de las dimensiones centrales del PC, permitiendo un avance progresivo sin generar una carga cognitiva excesiva. Esta configuración también resulta adecuada para su implementación en contextos educativos reales, donde se dispone de un tiempo limitado para la realización de actividades inmersivas.

4.2 Mecánicas de juego y aprendizaje

Los desafíos planteados en Revipec están diseñados para potenciar las habilidades clave del PC a través de la continua interacción con los objetos virtuales. Mediante actividades progresivas, los jugadores exploran conceptos fundamentales como el diseño de algoritmos, la abstracción y la evaluación de soluciones, esto promueve un aprendizaje basado en la exploración y el razonamiento lógico.

Cada una de las actividades ha sido cuidadosamente adaptada para asegurar una transición fluida entre el aprendizaje y la resolución de problemas. En la Tabla 1 se presentan los desafíos que componen Revipec, junto con las habilidades específicas del PC que se refuerzan en cada una de ellas. Estas actividades incluyen tareas como predecir trayectorias, planificar rutas con restricciones, identificar patrones, administrar recursos y organizar estructuras lógicas. Todas ellas están diseñadas para abordar distintas dimensiones del PC, como la descomposición de problemas, la abstracción, el pensamiento algorítmico, el reconocimiento de patrones y la evaluación de soluciones.

Tabla 1. Actividades de Revipec y habilidades reforzadas.

Actividad	Descripción	Habilidades Reforzadas
Trayectoria de Beabot	Predecir la ruta de un robot siguiendo un conjunto de reglas.	Descomposición, pensamiento algorítmico y reconocimiento de patrones.
Ruta con batería	Planificar un camino eficiente con restricción de movimientos.	Descomposición, abstracción, pensamiento algorítmico y evaluación de soluciones.
Organización del torneo	Ordenar las tarjetas en un árbol de resultados de un torneo.	Descomposición, abstracción, pensamiento algorítmico y reconocimiento de patrones.

Ataque a los secuaces	Identificar los patrones de movimiento para seleccionar el objeto correcto.	Descomposición, pensamiento algorítmico y reconocimiento de patrones.
Selección de batería	Identificar la estructura correcta.	Pensamiento algorítmico, reconocimiento de patrones y evaluación de
Construcción del puente	Administrar recursos para cruzar un camino de manera óptima.	Descomposición, abstracción, pensamiento algorítmico y evaluación de soluciones.
Carga de energía	Optimizar la capacidad de carga sin sobrepasar el límite	Descomposición, abstracción, pensamiento algorítmico y evaluación de soluciones

Para resolver las actividades presentadas, se hará uso de los Oculus Touch Controllers, permitiendo al usuario la manipulación de objetos, la selección de opciones y el desplazamiento dentro del entorno virtual. El sistema se ha diseñado para que la navegación y la interacción sean intuitivas, asegurando que los estudiantes puedan desenvolverse en la experiencia con pocos conocimientos previos en el uso de dispositivos de RV. Cada acción realizada por el jugador genera respuestas visuales y auditivas inmediatas, lo que refuerza la comprensión del entorno y promueve una retroalimentación constante.

De igual manera, se evitó la sobrecarga cognitiva mediante la limitación intencionada de elementos interactivos simultáneos, para mantener el enfoque del usuario en los aspectos clave de los retos presentados. El entorno proporciona un recorrido progresivo que ofrece una estructura de juego envolvente y motivadora. La combinación de mecánicas de gamificación con desafíos secuenciales refuerza el interés de los estudiantes y fomenta un aprendizaje significativo.

4.3 Aplicación en contextos educativos

Revipec está diseñado como un complemento para los contenidos curriculares que se abordan en las clases de pensamiento computacional, ofreciendo un espacio donde poner en práctica los conceptos aprendidos mediante la interacción en un entorno inmersivo. En lugar de ser un reemplazo para las sesiones de clase tradicionales, su propósito es actuar como un recurso adicional que facilite la consolidación de habilidades a través de experiencias activas y significativas.

Su diseño progresivo permite a los estudiantes avanzar de forma autónoma a través de las misiones, con una curva de dificultad ajustada que les guía desde los desafíos introductorios hasta las situaciones más complejas. Esta estructura no solo facilita la comprensión paulatina de los conceptos, sino que promueven la autonomía del estudiante al permitirle explorar y aprender a su propio ritmo dentro del entorno inmersivo.

Gracias a estas características, Revipec puede ser integrado como una herramienta didáctica versátil que favorece la participación activa y el aprendizaje significativo. Su implementación tiene el potencial de enriquecer la enseñanza del PC en contextos pedagógicos reales, ofreciendo una experiencia educativa innovadora acorde a las necesidades del siglo XXI.

4.4 Evaluación del aprendizaje

Para evaluar el impacto de Revipec en el desarrollo de habilidades del PC, se plantea la aplicación de un diseño cuasiexperimental que contemple la comparación entre un grupo experimental y un grupo de control. El grupo experimental contará con acceso a Revipec para complementar su aprendizaje, mientras que el grupo de control abordará el contenido mediante métodos tradicionales.

Ambos grupos serán sometidos a una prueba previa y una prueba posterior, con el objetivo de medir el progreso en las habilidades desarrolladas. Como instrumento de medición se empleará una adaptación del Beginner’s Computational Thinking Test (BCTt) [16], diseñado específicamente para estudiantes de primaria y secundaria. Esta prueba ha sido validada en diversos estudios y permite la evaluación de las dimensiones fundamentales del PC, sin requerir conocimientos de programación formal.

El uso del BCTt permitirá obtener datos cuantitativos sobre el aprendizaje logrado, brindando un punto de comparación entre ambos grupos y contribuyendo a la validación pedagógica de la propuesta.

5 Conclusiones

El desarrollo de Revipec representa una propuesta innovadora que complementa la enseñanza del Pensamiento Computacional (PC) integrando la Realidad Virtual (RV). La aplicación no solo ofrece una experiencia atractiva para los estudiantes, sino que refuerza las habilidades clave del PC.

A diferencia de enfoques más tradicionalistas, que generalmente dependen de plataformas digitales convencionales o actividades teóricas, Revipec brinda un aprendizaje experiencial, donde los estudiantes interactúan activamente con los desafíos estructurados en una narrativa envolvente. La implementación de la metodología MEDEERV ha permitido garantizar la alineación pedagógica del sistema, asegurando que cada interacción dentro del entorno virtual contribuya al desarrollo de habilidades cognitivas fundamentales.

Si bien la propuesta aún no ha sido evaluada empíricamente, el diseño estructurado y su fundamentación en marcos teóricos y metodológicos sólidos sientan una base confiable para evaluaciones futuras. Se tiene previsto el realizar pruebas controladas que permitan comparar el impacto de la herramienta frente a métodos tradicionales, utilizando instrumentos validados como el BCTt.

Los hallazgos derivados de estas pruebas permitirán generar nueva evidencia sobre el potencial de la RV en el ámbito educativo, además de establecer directrices para el diseño de futuros entornos inmersivos dirigidos al aprendizaje del PC.

Agradecimientos. El autor principal agradece a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo otorgado para el desarrollo de esta investigación. Asimismo, se reconoce el respaldo del Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX) de la Universidad Autónoma de Zacatecas (UAZ), cuyo acompañamiento académico y recursos tecnológicos hicieron posible la realización de este trabajo.

Referencias

1. J. M. Wing, "Computational thinking's influence on research and education for all Influenza del pensiero computazionale nella ricerca e nell'educazione per tutti", *Italian Journal of Educational Technology*, vol. 25, núm. 2, pp. 7–14, 2017, doi: 10.17471/24994324/922.
2. V. J. Shute, C. Sun, y J. AsbellClarke, "Demystifying computational thinking", el 1 de noviembre de 2017, Elsevier Ltd. doi: 10.1016/j.edurev.2017.09.003.
3. K. Harangus y Z. Kátai, "Computational thinking in secondary and higher education", en *Procedia Manufacturing*, Elsevier B.V., 2020, pp. 615–622. doi: 10.1016/j.promfg.2020.03.088.
4. N. Pellas, S. Mystakidis, y I. Kazanidis, "Immersive Virtual Reality in K12 and Higher Education: A systematic review of the last decade scientific literature", *Virtual Real*, vol. 25, núm. 3, pp. 835–861, 2021, doi:10.1007/s10055-020-00489-9.
5. D. Allcoat, T. Hatchard, F. Azmat, K. Stansfield, D. Watson, y A. von Mühlénen, "Education in the Digital Age: Learning Experience in Virtual and Mixed Realities", *Journal of Educational Computing Research*, vol. 59, núm. 5, pp. 795–816, ene. 2021, doi: 10.1177/0735633120985120.
6. G. A. Torres Samperio, A. Franco Arcega, M. D. J. Gutiérrez Sánchez, y A. Suarez Navarrete, "Metodología Para El Modelado De Sistemas De Realidad Virtual Para El Aprendizaje En Dispositivos Móviles", 2017. [En línea]. Disponible en: <http://itcelaya.edu.mx/ojs/index.php/pistas>
7. J. Wing, "Computational Thinking", *Commun ACM*, vol. 49, pp. 33–35, mar. 2006, doi: 10.1145/1118178.1118215.
8. R. Lege y E. Bonner, "Virtual reality in education: The promise, progress, and challenge", *JALT CALL Journal*, vol. 16, núm. 3, pp. 167–180, 2020, doi: 10.29140/jaltcall.v16n3.388.
9. Y. Xin, "Influence of Learning Engagement on Learning Effect under a Virtual Reality (VR) Environment", *International Journal of Emerging Technologies in Learning*, vol. 17, núm. 5, pp. 226–237, 2022, doi: 10.3991/ijet.v17i05.29451.
10. F. Joseph Agbo, S. Sunday Oyelere, J. Suhonen, y M. Tukiainen, "iThinkSmart: Immersive Virtual Reality Mini Games to Facilitate Students' Computational Thinking Skills", en *Proceedings of the 21st Koli Calling*

- International Conference on Computing Education Research, en Koli Calling '21. New York, NY, USA: Association for Computing Machinery, 2021. doi: 10.1145/3488042.3489963.
11. W. Zhao y V. J. Shute, "Can playing a video game foster computational thinking skills?", *Comput Educ*, vol. 141, nov. 2019, doi: 10.1016/j.compedu.2019.103633.
 12. M. Zapata-Caceres, E. Martin-Barroso, y M. Roman-Gonzalez, "Collaborative Game-Based Environment and Assessment Tool for Learning Computational Thinking in Primary School: A Case Study", *IEEE Transactions on Learning Technologies*, vol. 14, núm. 5, pp. 576–589, oct. 2021, doi: 10.1109/TLT.2021.3111108.
 13. J. Lee, C. Joswick, y K. Pole, "Classroom Play and Activities to Support Computational Thinking Development in Early Childhood", *Early Child Educ J*, vol. 51, núm. 3, pp. 457–468, mar. 2023, doi: 10.1007/s10643-02201319-0.
 14. M.-J. Tsai, J.-C. Liang, y C.-Y. Hsu, "The Computational Thinking Scale for Computer Literacy Education", *Journal of Educational Computing Research*, vol. 59, núm. 4, pp. 579–602, nov. 2020, doi: 10.1177/0735633120972356.
 15. H. Belmar, "Review on the teaching of programming and computational thinking in the world", 2022. doi: <https://doi.org/10.3389/fcomp.2022.997222>.
 16. M. Zapata-Caceres, E. Martin-Barroso, y M. Roman-Gonzalez, "Computational Thinking Test for Beginners: Design and Content Validation", en 2020 IEEE Global Engineering Education Conference (EDUCON), IEEE, abr. 2020, pp. 1905–1914. doi: 10.1109/EDUCON45650.2020.9125368.

Evaluating the Impact of Headset Weight on User Comfort in Immersive

Gabriel Alfonso Rodríguez Puerto ^[0009-0009-1934-0826], Nicolás Daniel García Sierra ^[0009-0007-4727-0891],
Daryi Valentina Vargas Umbarila ^[0009-0005-3215-5634], Carlos J. Latorre-Rojas ^[0000-0003-2417-7145],
Wilson J. Sarmiento ^[0000-0001-7903-8316]
Universidad Militar Nueva Granada
Grupo de Investigación en Multimedia -GIM-Bogotá D.C., Colombia
gim@unimilitar.edu.do

Abstract. The comfort of a device significantly influences its usability and the user experience, especially for 3D interaction and immersive devices worn on the body. Ergonomic design and weight are crucial for comfort for Head-Mounted Displays (HMDs), also known as virtual/mixed reality headsets. While commercial HMDs are often modified for research in immersive system interaction labs, the effect of these modifications on user comfort is seldom evaluated. This study examines a case where an HMD was modified for research, prompting an investigation into how to assess the impact of device weight on user comfort. A study with 32 participants in control and test groups indicated that subjective self-report questionnaires are insufficient for evaluating this variable, emphasizing the necessity of identifying biometric measures that can better reflect the user's condition.

Keywords: Virtual/Mixed Reality, Headsets, Comfort Assessment, Head Mounted Display.

1 Introduction

The search for better solutions to develop immersive experiences has led to a wide range of 3D immersion and interaction devices [14,13]. Many of these devices are wearable, some even classified as smart clothing, which means that users wear them directly on their bodies, such as gloves, vests, wristbands, and, of course, Head-Mounted Displays (HMDs) [14,13].

In the case of HMDs, now commonly referred to as virtual/mixed reality headsets, their development has evolved from heavy and complex devices that were difficult to use to commercial products available in department stores [6].

Many of these HMDs have originated in research laboratories, either as completely new developments or as modifications of commercial devices, to explore new forms of interaction and immersion [6,8]. In these developments, device functionality is often prioritized. At the same time, ergonomic design and user comfort are typically addressed in the final stages of commercial product design, which happens in only a few cases. However, many modifications to commercial HMDs alter their weight and ergonomics, affecting usability and the overall user experience. Despite this, very few laboratories conduct comfort evaluations on changes made to HMDs [6,8].

This study addresses the need to evaluate the comfort of a commercial head-mounted display (HMD) modification. It raises the question of how such an intervention might impact the comfort level of HMD. This paper presents a comparative evaluation of the comfort of a modified HMD, specifically analyzing the effects of added weight. The research was conducted with a sample of 32 participants divided into two contrasting groups. The results indicate that, in relatively small samples, subjective self-report questionnaires are not a reliable tool for this type of evaluation.

The remainder of this document is structured as follows. Section 2 presents a review of relevant work on comfort evaluation in wearable devices, focusing on HMD. Section 3 presents the materials and methods, describing user experimentation, subjective metrics used, and deep learning analysis by long short-term

memory networks. Section 4 discusses the results of the subjective evaluations and deep learning analyses found, and the last section presents the final remarks and future work discussion.

2 Related work

Comfort has been a primary concern in the development of wearable devices. Regarding subjective metrics, perhaps one of the most pertinent contributions has been the proposal by Knight and Baber: The Comfort Rating Scales (CRS). As described, these scales measure comfort across six dimensions: emotion, attachment, harm, perceived change, movement, and anxiety. The CRS has been shown to reliably assess comfort in wearable computers, aiding designers in identifying areas for improvement [4].

Concerning Head-Mounted Displays (HMDs), comfort and usability evaluations have predominantly targeted highly specific issues, such as visual comfort [11] and the phenomenon of cybersickness [1,12]. Nevertheless, some literature reviews have highlighted the relevance of multiple factors—including weight and head-fixation ergonomics—impacting user satisfaction during device usage [3,9]. Indeed, certain studies have directly investigated the impact of device weight on the user's body and posture [7].

3 Materials and methods

3.1 Experimental Design

This study aimed to identify and validate strategies that enable the differentiation of discomfort engendered by the increased mass of a Head-Mounted Display (HMD). Consequently, the hypothesis for empirical verification was formulated, which is as follows:

***H1:** Self-report questionnaires designed to measure usability, immersion, or comfort dimensions can differentiate the impact of added weight on using an HMD*

A traditional experimental design contrasting two conditions was employed to test this hypothesis. For this experiment, the population was divided into two groups. The first group used an HMD modified with an A traditional experimental design contrasting two conditions and was employed to test this hypothesis. For this experiment, the population was divided into two groups. The first group used an HMD modified with an Insta360 X3 camera, which weighed 180 grams W_0 . The second group used an HMD modified with a dummy device weighing approximately 3 W_0 equivalent to 533 grams. The dummy was designed to replicate the approximate dimensions of the 360° camera but with significantly greater weight. It is imperative to elucidate that the baseline mass of the HMD is 722 grams. Consequently, the cohort comprising the first condition (camera) operated an HMD with a total mass of 900 grams, while the second condition (dummy) utilized an HMD exhibiting a total mass of 1253 grams. Fig. 1 shows two users in the test, (a) is a camera condition user, and (b) is a dummy condition user.

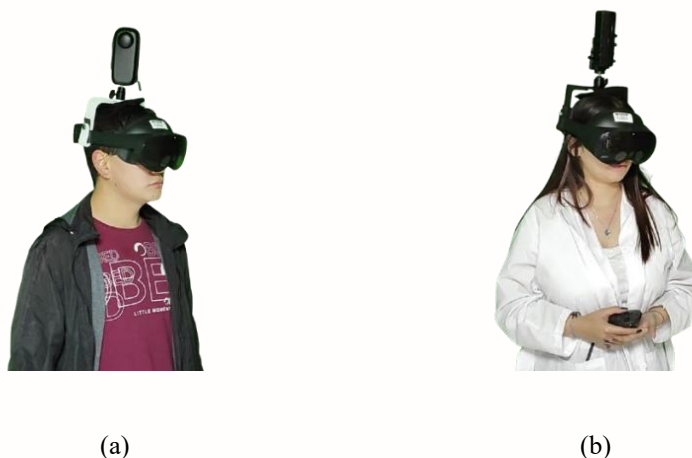


Fig. 1. In (a), a user wearing the HMD with the 360° camera, (b) shows a second user wearing the HMD with the dummy. It should be highlighted that the dimensions of the dummy device and the 360° camera are the same.

A 360° immersive video observation activity was selected as an interactive task. The main idea was to require users to move their bodies in all directions to explore the scene. To this end, two videos were recorded where a volleyball team performed passes in a circular formation (Fig. 2). The camera was positioned in the center of the formation, so the users had to rotate 360° to follow the ball's trajectory. Women's and men's teams were recorded at a resolution of 3840×1920 pixels and a frame rate of 29.97 fps. The videos were edited in post-production for the same duration: 1 minute and 30 seconds.

An event-counting task related to the volleyball team's training sessions was designed to encourage user movement. To minimize task-related bias, half of the participants in each experimental condition (camera and dummy) viewed the women's team video, while the other half viewed the men's team video. This approach ensured that any conclusions drawn would not depend on which video the participants watched.



Fig. 2. Panoramic (360°) videos recorded on experimental test. (a) women's volleyball team, and (b) men's volleyball team. The camera was placed in the center of the action.

3.2 Procedure

The main steps of the experimental procedure are described below.

Preliminary Preparation. This phase of the experimental procedure had two objectives. The first was to provide participants with a general overview of the study, without disclosing information that could introduce

bias, and, if necessary, to allow them to withdraw voluntarily. The second objective was to collect demographic and characterization data from participants. This phase followed the following steps:

- A welcome greeting was conducted, during which participants were informed that their participation was voluntary.
- Users completed a digital informed consent form, including authorization to use images (photographs and videos) for academic and scientific dissemination.
- The users filled out a demographic and characterization questionnaire.

Pretest. This phase aimed to familiarize the users with the HMD-based video viewing interaction. A simplified version of the main task was presented to ensure that users understood the procedure and any doubts were addressed as needed. The steps in this phase were the following:

- The participant read a short description explaining that they would view a 360° video showing a couple rowing at sea. They were instructed to observe how many sailboats appeared in the video. The video had a duration of 29 seconds.
- The users wore HMD and watched the video.
- After viewing, the users removed the headset and answered the corresponding question.
- Users were asked if they had any doubts about how to observe the immersive video or about the task itself. If concerns were raised, all necessary clarifications were provided.

Test. This was the core phase of the study. The participant completed the interactive task to be evaluated: watching a 360° video using a virtual reality headset. It was essential that the users clearly understood the interactive process to avoid introducing bias in the data analysis. This phase was conducted as follows:

- Once again, the participant read an instructional text. This time, they were informed that they would watch a 360° video in which a volleyball team performed passes in a circular formation. They were instructed to count how many times the ball fell to the ground and how many overhead passes (finger passes) were made.
- The participant wore the HMD and viewed the video.
- After the visualization task, the participant removed the headset and completed self-report questionnaires.

3.3 Population

For the present study, an invitation to voluntary participation was extended to the academic community. A total of 31 individuals participated in the experimental protocol, with one volunteer being excluded due to considerations of study symmetry. The experimental procedure was carried out with male 18 self-identified, female 13 self-identified, and a participant 1 who did not choose to disclose their gender identity. The mean age of the participant cohort was 20.8 years, with a standard deviation of 6.23 years. Regarding ocular conditions, 17 participants reported no diagnosed ophthalmological problems, while 15 presented with myopia, 12 with astigmatism, 1 with glaucoma, and 1 with lenticonus. It is pertinent to note that these conditions are not mutually exclusive. In addition, participants were asked about their experience with virtual reality technology. 17 respondents indicated no previous exposure to this technology, 11 reported previous interaction with an HMD, and 4 stated regular use of this technology.

3.4 Hardware, instrument, and data recollection

During the experimental trials, participants utilized the *Meta Quest Pro VR HMD*, characterized by 12 GB of RAM, 256 GB of internal storage, and powered by a Qualcomm Snapdragon XR2+ processor. The visual system incorporates two LCD panels with a resolution of 1800×1920 pixels per eye, providing a field of view (FOV) of 106° along the horizontal axis and 96° along the vertical axis. In this way, the baseline mass of the system is 722 grams without modifications. This HMD runs the *Meta Horizon OS V37* operative system, which is compatible with *Android 10*.

As indicated previously, after completion of the experimental protocol, all participants completed a battery of self-report questionnaires. The initial instrument administered was the Comfort Rating Scales (CRS), proposed by Knight et al. [4]. This psychometric tool facilitates the assessment of a user's subjective comfort level while wearing a body-borne device across six distinct dimensions, namely: Emotions (encompassing feelings, concerns regarding appearance, and relaxation); Attachment (about the proprioceptive experience of the device on the user's physique); Harm (reflecting the user's perception of potential physical effects or detriment to their corporeality); Perceived Change (addressing concerns about the user's subjective sensation of physical alteration or perturbation); Movement (relating to the device's influence on the user's physical locomotion); and Anxiety (denoting apprehension concerning the device's functionality, safety, and dependability) [4].

The second questionnaire aimed to evaluate cybersickness induced by visualization activity; specifically, the Cybersickness in Virtual Reality Questionnaire (CSQ-VR), as formulated by Kourtesis et al. [5], was applied. This instrument assesses three dimensions, Nausea, Vestibular, and Oculomotor, with two questions assigned to each dimension, each item carrying the same weight [5]. In addition, a user experience questionnaire, the Short User Experience Questionnaire (UEQ-S), was administered, evaluating its eight responses on two dimensions: hedonic and pragmatic quality [10]. Finally, the System Usability Scale (SUS) was used to assess perceived usability [2].

4 Data analysis and results

To evaluate the hypothesis, denoted as **H1**, a statistical contrast analysis was conducted on the response data obtained from the two experimental groups (Camera and Dummy). The initial step involved a Shapiro-Wilk normality test, which indicated that none of the response variables conformed to a normal distribution. Consequently, and in agreement with these findings, a Wilcoxon Rank-Sum test was performed for mean comparisons for each self-report questionnaire.

Table 1 shows the results of these analyzes, revealing that the results are nonconclusive, since the p-value in all contrast tests significantly exceeds the alpha level of 0.01. Therefore, for the present experiment, **H1** was rejected. These results may be attributable to two primary factors. Firstly, the limited number of test participants, given that the experiment was conducted with 16 participants per condition. A second potential contributing factor is the duration of the interaction, a mere 1 minute and 30 seconds, which may be an insufficient temporal interval for users to perceive the added mass of the HMD in a statistically significant manner. Despite the statistical result, a graphical analysis of the resultant data facilitates a more comprehensive interpretation.

Table 1. Summary of the results of statistical contrasts of user answers of the self-report questionnaire. In the last column, the p-value of the Wilcoxon Rank-Sum test.

Self-report Dimension		Camera		Dummy		p-value
		μ	σ	μ	σ	
CRS	Emotion	0.770	0.196	0.715	0.259	0.589
	Attachment	0.354	0.240	0.375	0.300	0.908
	Harm	0.972	0.049	0.875	0.246	0.460
	Perceived Change	0.583	0.267	0.597	0.214	0.848
	Movement	0.701	0.245	0.722	0.281	0.773
	Anxiety	0.881	0.187	0.687	0.363	0.262
CSR-VR	Nausea	0.281	0.204	0.200	0.065	0.128
	Vestibular	0.316	0.132	0.361	0.190	0.647
	Oculomotor	0.290	0.167	0.200	0.098	0.084
UEQ-S	Pragmatic	0.409	0.044	0.402	0.050	0.610
	Hedonic	0.475	0.041	0.448	0.063	0.234
SUS		0.798	0.109	0.820	0.126	0.449

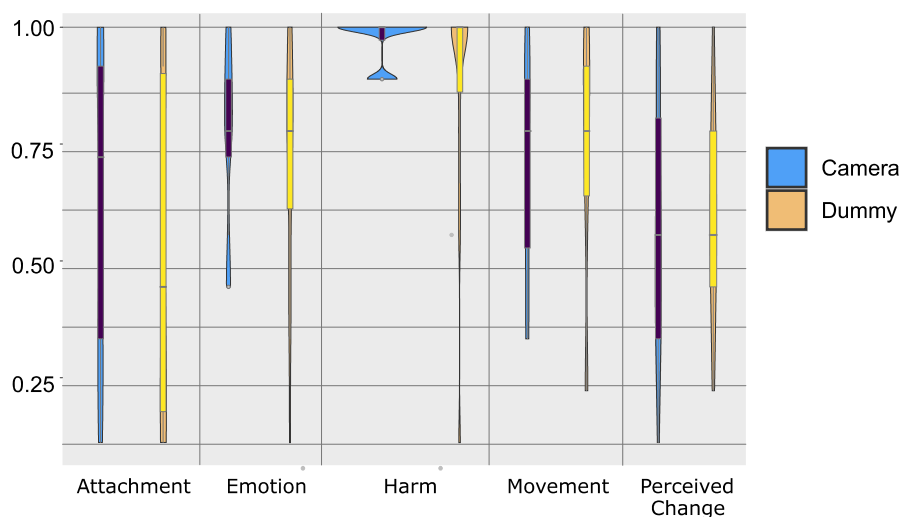


Fig. 3. Violin plot graphs that comparison the six dimensions of CRS questionnaire.

Fig. 3 shows a violin plot comparing the CRS dimensions under experimental conditions. Observationally, the dimensions of Attachment, Emotion, Movement, and Perceived Change exhibit similar behavioral patterns for both conditions, characterized by substantial data dispersion.

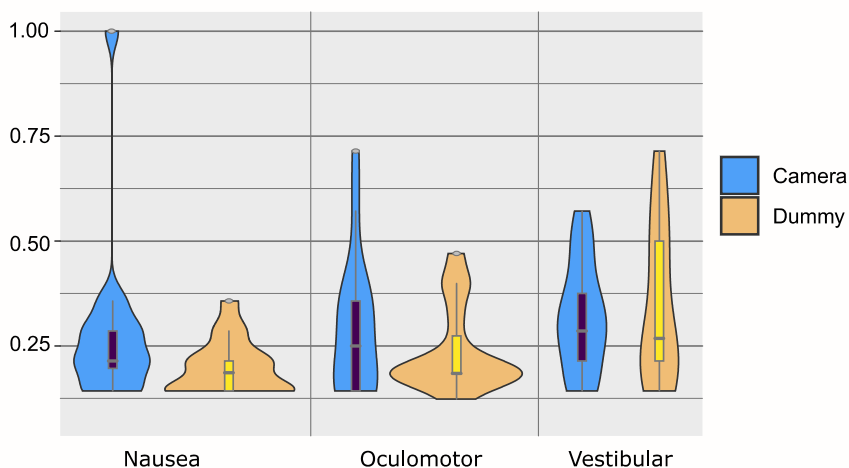


Fig. 4. Violin plot graphs that comparison the six dimensions of CRS-VR questionnaire.

In contrast, the Harm dimension, related to the perceived physical safety of the device, demonstrates discernible disparities in its probability distributions, where users reported that the HMD augmented with the camera was subjectively safer than the HMD integrated with the dummy weight. Regarding cybersickness, the violin graphs in Fig. 4 reveal that although statistically significant differences in mean values are absent, their respective probability distributions manifest divergent characteristics. A similar analytical approach can be applied to graphical representations of usability and user experience, Fig. 5.

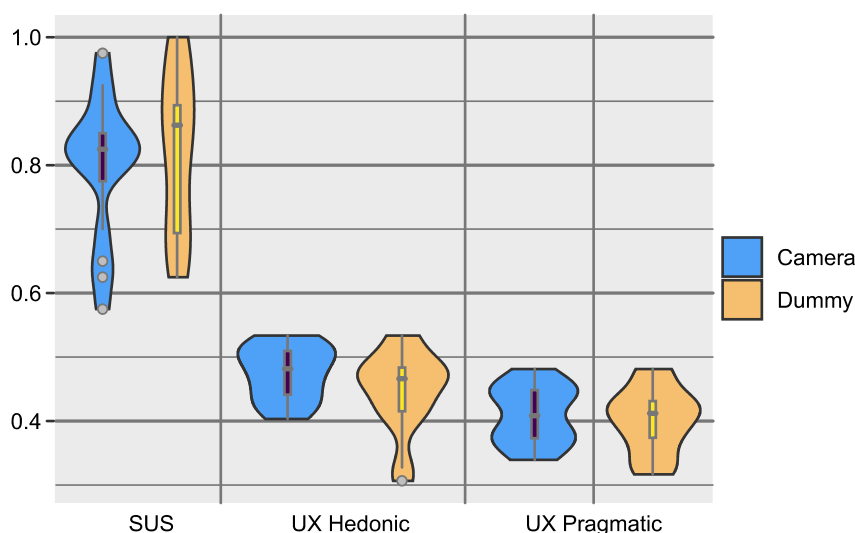


Fig. 5. Violin plot graphs that comparison the Hedonic and Pragmatic dimensions of UEQ-S, and SUS.

As mentioned earlier, the observations suggest the presence of inherent data heterogeneity, which implies that the differential weight between the two experimental conditions influenced user perception concerning motion sickness sensation and aspects of usability and user experience.

5 Final remarks and future work

This study compared two user groups interacting with a modified Head-Mounted Display (HMD) featuring added weight while performing a basic panoramic video visualization task. The study was conducted with a sample of 32 participants, divided into two contrast groups of equal size, comprising 16 users each. The statistical contrast results from the administered self-report questionnaires did not reveal statistically significant differences. Consequently, it cannot be affirmed that these questionnaires constitute an adequate instrument for assessing the impact of adding weight to an HMD under these conditions.

This research raises important questions. Could self-report questionnaires yield significant differences with a larger user cohort and extended interaction duration? For this reason, future work will perform tests with the largest user set and richest interaction involving other tasks such as selection, manipulation, navigation, and consequently, the longest interaction time. Also, biometric data must be included to look for ways to identify differences in user behavior caused by weight modification in an HMD.

Acknowledgements. The authors thank the Universidad Militar Nueva Granada (UMNG) for supporting this research endeavor. The authors, Gabriel Rodríguez and Wilson J. Sarmiento, faculty members (occasional and one full-time) at UMNG, are acknowledged for the time dedicated to this research. The authors, Valentina, Nicolás, and Carlos, gratefully acknowledge the financial assistance provided by the university's welfare and graduate student funds.

References

1. Ang, S., Quarles, J.: Reduction of cybersickness in head mounted displays use: A systematic review and taxonomy of current strategies. *Frontiers in Virtual Reality* Volume 4 - 2023 (2023). <https://doi.org/10.3389/frvir.2023.1027552>.
2. Brooke, J.: SUS-A quick and dirty usability scale, pp. 252–261. CRC Press, LPatrick W. Jordan and B. Thomas and Ian Lyall McClelland and Bernard Weerdmeester, 1rd edn. (pages = 1996,). <https://doi.org/10.1201/9781498710411>.

3. Cheng, D., Wang, Q., Liu, Y., Chen, H., Ni, D., Wang, X., Yao, C., Hou, Q., Hou, W., Luo, G., Wang, Y.: Design and manufacture of head-mounted displays: A review and outlook. *Light: Advanced Manufacturing* 2(LAM2020120035), 350 (2021). <https://doi.org/10.37188/lam.2021.024>.
4. Knight, J.F., Baber, C.: A tool to assess the comfort of wearable computers. *Human Factors* 47(1), 77–91 (2005). <https://doi.org/10.1518/0018720053653875>.
5. Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., MacPherson, S.E.: Cybersickness in virtual reality questionnaire (CSQ-VR): A validation and comparison against ssq and vrsq. *Virtual Worlds* 2(1), 16–35 (2023). <https://doi.org/10.3390/virtualworlds2010002>.
6. Kress, B., Starner, T.: A review of head-mounted displays (HMD) technologies and applications for consumer electronics. In: Kazemi, A.A., Kress, B.C., Thibault, S. (eds.) *Photonic Applications for Aerospace, Commercial, and Harsh Environments IV*. vol. 8720, p. 87200A. International Society for Optics and Photonics, SPIE (2013). <https://doi.org/10.1117/12.2015654>.
7. Lee, Y., Park, D., Kim, Y.M.: The effect of wearing a head-mounted display on the boundaries of the cervical range of motion based on perceived comfort in a static posture. *Virtual Reality* 27(2), 815–828 (2023). <https://doi.org/10.1007/s10055-022-00684-w>.
8. Lu, Y., Ota, K., Dong, M.: An empirical study of vr head-mounted displays based on vr games reviews. *ACM Games* 2(3) (Aug 2024). <https://doi.org/10.1145/3665988>.
9. Renganayagalu, S.K., Mallam, S.C., Nazir, S.: Effectiveness of vr head mounted displays in professional training: A systematic review. *Technology, Knowledge and Learning* pp. 1–43 (2021). <https://doi.org/10.1007/s10758-020-09489-9>.
10. Schrepp, M., Thomaschewski, J., Hinderks, A.: Design and evaluation of a short version of the user experience questionnaire (ueq-s). *International Journal of Interactive Multimedia and Artificial Intelligence* 4(6), 103–108 (12/2017 2017). <https://doi.org/10.9781/ijimai.2017.09.001>.
11. Shrestha, P.K., Pryn, M.J., Jia, J., Chen, J.S., Fructuoso, H.N., Boev, A., Zhang, Q., Chu, D.: Accommodation-free head mounted display with comfortable 3d perception and an enlarged eye-box. *Research* (2019). <https://doi.org/10.1109/TVCG.2015.2473855>.
12. de Souza, J., Tartz, R.: Visual perception and user satisfaction in video see-through head-mounted displays: a mixed-methods evaluation. *Frontiers in Virtual Reality* Volume 5 - 2024 (2024). <https://doi.org/10.3389/frvir.2024.1368721>.
13. Spittle, B., Frutos-Pascual, M., Creed, C., Williams, I.: A review of interaction techniques for immersive environments. *IEEE Transactions on Visualization and Computer Graphics* 29(9), 3900–3921 (2023). <https://doi.org/10.1109/TVCG.2022.3174805>.
14. Suh, A., Prophet, J.: The state of immersive technology research: A literature analysis. *Computers in Human Behavior* 86, 77–90 (2018). <https://doi.org/10.1016/j.chb.2018.04.019>.

Sistema Inteligente de Atendimento à Emergências: Viabilizando Gêmeo Digital e Otimização de Recursos em Cidades Inteligentes

Vinicius Alves Marques¹, Leandro Pupo Natale^{1,2}

¹ Faculdade de Computação e Informática - Universidade Presbiteriana Mackenzie, São Paulo, SP, Brasil.

² Empresa de Processamento de Dados do Município de São Paulo, São Paulo, SP, Brasil.

Abstract. Emergency transportation is essential in urban services, especially in high-demand scenarios. The integration of the Internet of Things (IoT) can revolutionize emergency response efficiency. This project proposes the use of IoT to monitor and optimize the operation of emergency vehicle fleets. Sensors and connected devices will enable real-time tracking and the transmission of critical patient data to hospitals. This will lower response times and enhance patient safety. The research will analyze technical feasibility and implementation challenges. The goal is to improve emergency transportation in smart cities.

Keywords: Transporte de Emergência Inteligente, IoT, Gêmeo Digital Urbano.

1 Introdução

Uma cidade inteligente é definida como um espaço urbano que utiliza Tecnologias de Informação e Comunicação (TIC) para melhorar a qualidade de vida de seus cidadãos, otimizar a gestão de recursos, promover o crescimento econômico e garantir a sustentabilidade [6][7][9]. Para a criação e manutenção de cidades inteligentes, é necessário investimento em capital humano e social, estabelecer governança participativa, gestão inteligente de recursos e mobilidade urbana eficiente, além de todos estes aspectos garantirem a segurança e privacidade dos munícipes.

De forma adicional, um gêmeo digital é uma representação virtual de uma entidade física ou sistema, que espelha seus atributos e comportamentos em tempo real [8]. Analisando o contexto urbano, as tecnologias empregadas para tornar uma cidade inteligente podem também serem usadas para o estabelecimento de um gêmeo digital possibilitando além da visualização das informações em tempo real, a simulação e otimização de diferentes aspectos da vida urbana [6][7][9].

Neste sentido, o monitoramento de ambulâncias e veículos de emergência emerge como um componente essencial, permitindo assim uma resposta rápida às emergências ocorridas no contexto urbano, proporcionando melhorias no atendimento, e gestão de recursos de forma mais eficiente.

A adoção de tecnologias é um dos pilares das cidades inteligentes, permitindo o monitoramento e a automação de serviços urbanos, bem como o compartilhamento de informações em tempo real com administrações municipais, prestadores de serviços e residentes. As cidades inteligentes também adotam tecnologias cognitivas para construir resiliência, melhorar a qualidade de vida e otimizar o uso da infraestrutura.

Ademais, avanços tecnológicos, como o da Internet das Coisas (IoT), computação em nuvem, Inteligência Artificial (IA) e análise de dados passaram a ser amplamente aplicadas em diversos setores urbanos, desde o transporte e energia até saúde e a segurança pública.

Neste sentido, este artigo apresenta uma proposta de solução para a otimização do transporte de emergência em cidades inteligentes, a partir da integração de tecnologias de IoT e da implementação de uma arquitetura de software integrador orientada a objetivos. O projeto está sendo desenvolvido em colaboração com a PRODAM-SP (Empresa de Tecnologia da Informação e Comunicação do Município de São Paulo) e com a Prefeitura de São Paulo, no âmbito de um acordo de cooperação institucional com a Universidade Presbiteriana Mackenzie. Esta iniciativa integra um projeto maior voltado à criação de um gêmeo digital da cidade, fundamentado na estratégia de living labs, com o objetivo de promover inovação aberta, desenvolvimento tecnológico e melhoria contínua dos serviços urbanos.

A proposta aqui discutida busca promover a interoperabilidade entre diferentes sistemas e dispositivos, com ênfase no monitoramento e controle eficaz da frota de emergência, contribuindo para respostas mais rápidas e coordenadas em situações críticas. Desta forma, a solução proposta para a cidade inclui a utilização de forma integrada de dispositivos médicos conectados no interior do veículo, dispositivos técnicos de coleta de dados do veículo, como velocidade, aceleração, etc enviando todo este conjunto de dados para que uma central de comando tenha informações detalhadas para a tomada de decisão em momentos críticos.

A arquitetura de software proposta facilita a integração de dados provenientes de diversas fontes, incluindo:

- Um aplicativo móvel para o líder socorrista, que se conecta a um dispositivo On-Board Diagnostics II (OBDII) instalado no veículo de emergência.
- Informações do veículo (nível de combustível, aceleração etc.).
- Dados de georreferenciamento do celular.
- Dispositivos médicos conectados no interior do veículo (estetoscópio, oxímetro etc.).

A integração com o backend da solução proposta possibilita o armazenamento, a apresentação de dados em diferentes formatos para dashboards de análise e a criação de um gêmeo digital da cidade no contexto de transportes de emergência. Tais funcionalidades visam garantir uma operação mais coordenada e eficiente dos serviços de emergência, aprimorando a resposta a incidentes críticos.

Adicionalmente aspectos relacionados a Interação Humano-Computador (IHC) estão sendo consideradas desde a concepção do projeto garantindo que a solução atenda de forma unificada as necessidades técnicas de integração dos dispositivos conectados e as situações sensíveis com o transporte de emergência. Desta forma foi definido como elemento central do projeto o uso de um aplicativo mobile devido não apenas à sua praticidade técnica mas também em sua capacidade de integrar diferentes fontes de dados.

Esta opção oferece uma experiência de usuários simplificada a ponto de quase se tornar ubíqua, em observância as questões específicas de um socorrista o qual necessita de agilidade durante a execução de seu trabalho.

2 Trabalhos relacionados

Com o advento de sistemas mais conectados e inteligentes, a possibilidade de aprimorar os serviços essenciais de emergência através da tecnologia IoT torna-se viável e essencial para a sustentabilidade e eficácia dos serviços urbanos. Esta seção delinea metas específicas para o monitoramento e controle eficiente da frota de emergência, desenvolvimento de aplicações informativas para gestores e cidadãos, e a crucial transmissão de dados médicos em tempo real. Estes objetivos são projetados para garantir uma operação mais coordenada e eficiente dos serviços de emergência, melhorando significativamente a resposta a incidentes críticos e, por consequência, salvando mais vidas [5][10].

Um exemplo prático é o apresentado por [4], o qual propõe uma solução abrangente para o monitoramento de veículos de emergência e a detecção de acidentes. A proposta inclui a instalação de diferentes sensores em veículos de emergência comuns, permitindo a identificação automática de acidentes. Esses sensores conectados à internet tem a capacidade de enviar informações detalhadas sobre o ocorrido, possibilitando a classificação dos acidentes e auxiliando no despacho adequado destes veículos para os locais das ocorrências. O projeto também incorpora algoritmos de roteirização para determinar a melhor rota, evitando ruas com trânsito lento e outros incidentes. São utilizados como sensores o módulo NEO-6M GPS e o módulo SIM900A-GSM, e como microprocessadores, o Raspberry Pi.

O estudo de [1] propõem um sistema de ambulâncias inteligentes as quais equipadas com dispositivos que possibilitam o georreferenciamento para monitorar o posicionamento das ambulâncias na cidade. Este sistema propõe o envio de dados a uma central de controle e exibidos por meio de um dashboard em tempo real. Os administradores do sistema podem então, selecionar a ambulância mais bem localizada para o atendimento. Esse sistema funciona em conjunto com a gestão hospitalar, integrando diferentes hospitais, para identificar aquele com capacidade ociosa para atendimento de emergência.

A solução proposta por [11] também aborda o georreferenciamento das ambulâncias para otimizar e facilitar a definição do veículo a ser enviado para atender os pacientes. No entanto, eles incluem o uso de dispositivos responsáveis por monitorar os sinais vitais dos pacientes, como batimentos cardíacos e oxigenação do sangue. Esses dispositivos medem os sinais vitais em tempo real e enviam os dados para o hospital que receberá o paciente, permitindo que a equipe médica esteja informada antecipadamente e possa agilizar o tratamento.

A proposta de [12], no contexto da cidade de Manchester, Reino Unido, avalia o contexto de grandes desastres como terremotos, tsunamis, etc em que principalmente indivíduos vulneráveis como crianças e idosos estão mais propensos a ficarem presos em tais situações. A proposta de incluir equipamentos de IoT em ambulâncias visam aprimorar sua capacidade operacional de forma a promover dados a serem considerados no roteamento, otimizando o deslocamento.

Considerando um gêmeo digital urbano como uma estrutura virtual que espelha, em tempo real, os elementos e comportamentos de um sistema físico, sua aplicação no contexto do transporte de emergência permite avanços significativos na resposta a incidentes críticos. A integração dos dados clínicos do paciente, coletados por dispositivos vestíveis embarcados na ambulância, com os dados técnicos operacionais do veículo (como velocidade, aceleração, localização e status dos sistemas), viabiliza o monitoramento contínuo e a detecção precoce de alterações no estado de saúde.

O histórico consolidado dessas informações permite a construção de réplicas digitais dinâmicas, capazes de fornecer uma visão preditiva e analítica sobre o atendimento pré-hospitalar. Tecnologias como o Enhanced Intelligent Connectivity Framework (E-ICF) [13] demonstram como o espelhamento digital do ambiente físico pode viabilizar simulações e análises em tempo real, antecipando falhas operacionais, sugerindo manutenções preditivas, e oferecendo recomendações personalizadas de saúde durante o atendimento, promovendo maior segurança e eficiência nos serviços urbanos de emergência.

3 Proposta da arquitetura da solução integração

Com base nos objetivos definidos, este projeto propõe uma solução integrada para o monitoramento e controle de dados veiculares, correlacionando informações da frota de ambulâncias com os sinais vitais dos pacientes transportados. A partir da captura de dados em tempo real, provenientes de dispositivos conectados, como estetoscópios eletrônicos, é possível acompanhar a condição do paciente durante o trajeto, e gerar um histórico detalhado do atendimento, contemplando diferentes momentos: antes, durante e após o transporte de urgência.

A arquitetura do sistema foi projetada em camadas, demonstrada na Figura 1, garantindo modularidade, escalabilidade e integração eficiente entre dispositivos e aplicações. A camada de apresentação suporta tanto aplicações móveis quanto web, permitindo acesso a informações em tempo real por diferentes perfis de usuários, como gestores, socorristas e profissionais de saúde. Já a camada de integração é composta por um barramento de serviços baseado em microsserviços, viabilizando maior granularidade no desenvolvimento, maior flexibilidade para evolução da aplicação e crescimento horizontal eficiente em ambientes Kubernetes. Para armazenamento, a camada de dados utiliza o banco NoSQL MongoDB, proporcionando eficiência na persistência e extração de dados, especialmente no contexto de monitoramento contínuo.

A arquitetura do projeto está desenvolvida com base em requisitos de alta disponibilidade, escalabilidade, interoperabilidade fundamentais em sistemas de gerenciamento de transporte de emergência. Estes requisitos não funcionais são endereçados através de uma abordagem baseada em microsserviços orquestrada via Kubernetes. A utilização de um banco de dados NoSQL como o MongoDB permite lidar com o grande volume de dados semiestruturados estimados em uma operação crítica para uma cidade como São Paulo.

Cabe destacar a inserção deste projeto como uma fonte geradora de dados para a criação de um gêmeo digital urbano, responsável por produzir uma das camadas do contexto de emergências médicas. Diferente de soluções, como as citadas anteriormente que operam como um simples painel de monitoramento, este projeto está sendo modelado de forma funcional representando o fluxo de atuação (mobilização, atendimento, transporte e chegada), sincronização do ambiente físicos e suas réplicas digitais, disponibilização de dados históricos de forma a capacitar a simulação, predição de eventos e avaliação de cenários de resposta.

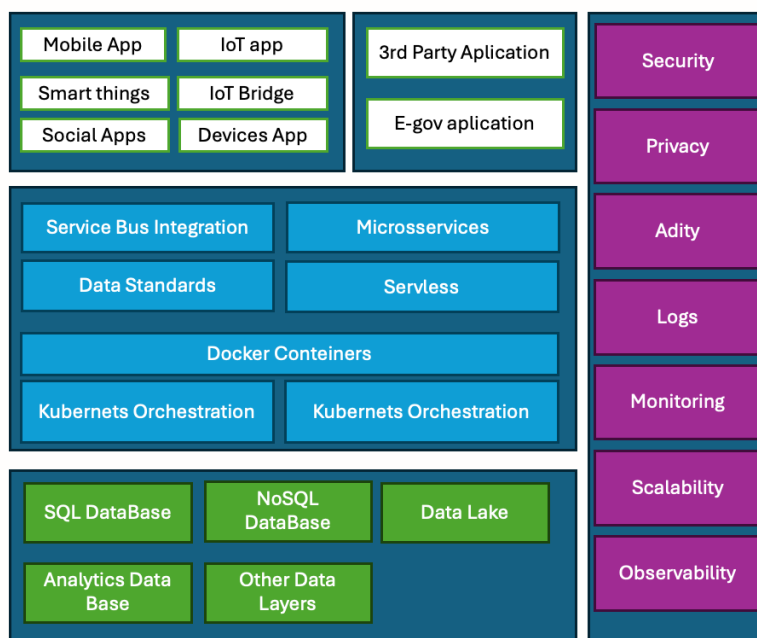


Fig. 1. Arquitetura de integração em camadas (termos em inglês mantendo termos originais das tecnologias).

Essa abordagem amplia o escopo do sistema para além do monitoramento e se posiciona como um instrumento de apoio à decisão estratégica em tempo real, promovendo não apenas visibilidade, mas também inteligência operacional para os serviços públicos de saúde.

4 Apresentação da aplicação proposta

O sistema de monitoramento de veículos de emergência segue as definições da arquitetura em camadas implementadas de forma a permitir a integração entre os dispositivos móveis e o backend da aplicação, conforme apresentado no diagrama de arquitetura da figura 2.

A solução proposta está composta por um aplicativo intuitivo, projetado para prover tanto a interação dos responsáveis pela equipe de socorristas quanto o monitoramento em tempo real e o controle eficiente da frota de ambulâncias, aproveitando a tecnologia GPS embarcada no dispositivo móvel, melhor detalhado no próximo item e apresentação na figura 3 da versão inicial do aplicativo. Ao permitir o acompanhamento constante da localização e do status de cada veículo, o sistema possibilita uma tomada de decisões mais ágil e informada em situações de emergência.

Esta definição de utilização do dispositivo móvel como centralizador dos dados de todos os dispositivos do veículo e de localização se dá pelas dificuldades encontradas em cidades de regiões de sombra de sinal, e facilidade de gerenciamento da conectividade com as redes 4G e 5G das operadoras.

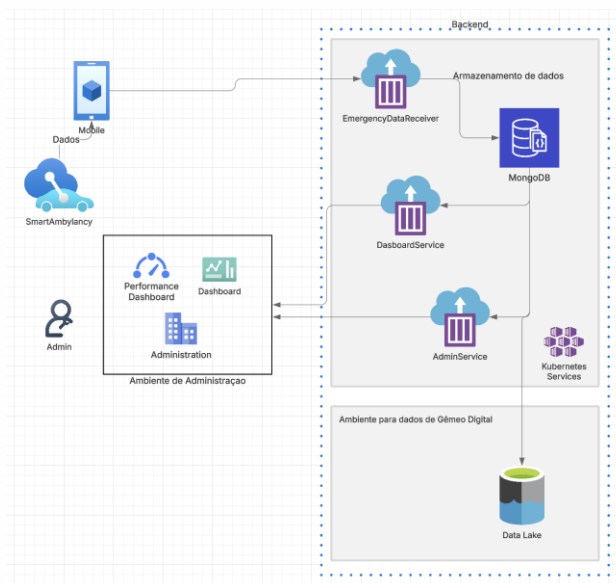


Fig. 2. Visão de implementação da solução proposta.

Desta forma, o responsável pela equipe socorrista deve realizar a autenticação no aplicativo, conforme apresentado na figura 3 (a), em seguida selecionar o veículo que será utilizado, conforme figura 3 (b), e ainda ter disponível os dados coletados do dispositivo OBDII, como apresentado na figura 3 (c).

Finalmente Através dos dados coletados por este dispositivo e enviados ao backend possibilitará aos gestores das frotas de emergência a coordenação e o despacho eficiente dos veículos otimizando o uso dos recursos e garantindo a eficiência operacional. Com a coleta dos quilômetros percorridos, consumo de combustível, média de velocidade e outros dados permitirá a análise e identificação de manutenção veicular, otimização de rotas, minimizando o tempo de inatividade e garantindo a segurança.

Conforme apresentado, este aplicativo mobile funcionará como um centralizador do monitoramento e será responsável pela transmissão em tempo real dos sinais vitais das vítimas para o hospital de destino, utilizando a tecnologia IoT. Ao prover dados precisos e atualizados sobre o estado da vítima durante o transporte, o sistema permite que a equipe médica se prepare antecipadamente para o atendimento imediato, otimizando a resposta médica e aumentando as chances de recuperação. A seguir estão listadas as funcionalidades necessárias relacionadas à coleta dos dados de sinais vitais do paciente em atendimento.

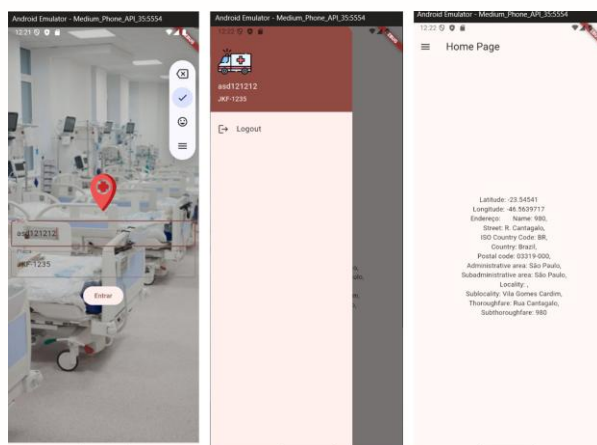


Fig. 3. Capturas de tela da versão inicial da aplicação mobile

- **Monitoramento em Tempo Real:** Sensores conectados à vítima coletam e transmitem continuamente os sinais vitais (frequência cardíaca, pressão arterial, temperatura, oximetria, ECG, entre outros) para o hospital.
- **Transmissão de Dados:** O sistema garante a transmissão segura e confiável dos dados, mesmo em áreas com cobertura de rede limitada.
- **Troca de Informações:** Além dos sinais vitais, o sistema permite a troca de informações sobre o estado geral da vítima, incluindo histórico médico, alergias, medicações em uso e outros dados relevantes.
- **Coordenação:** A plataforma facilita a comunicação e a coordenação entre as equipes de resgate e os profissionais de saúde no hospital, garantindo uma transição suave e eficiente no atendimento.
- **Preparação Antecipada:** Ao receber os dados da vítima antecipadamente, a equipe médica pode se preparar para o atendimento imediato, definindo prioridades, mobilizando recursos e preparando equipamentos.
- **Otimização da Resposta Médica:** A combinação do monitoramento em tempo real, da troca de informações e da preparação antecipada resulta em uma resposta médica mais rápida, eficiente e personalizada, aumentando as chances de recuperação da vítima

4.1 Painel para Gestão de Informações de Frota

Está previsto para este projeto uma aplicação front-end responsável pela exibição do gêmeo digital da frota de emergência da cidade. Estes dados estão previstos para serem exibidos através de painéis especificamente direcionado aos gestores municipais, com objetivo de fornecer uma plataforma centralizada para a visualização de dados históricos e o controle em tempo real da frota de emergência. Esta plataforma poderá oferecer acesso rápido e intuitivo a dados atualizados, permitindo uma tomada de decisões mais informada e transparente nas operações de emergência.

Estes painéis de visualização personalizáveis poderão oferecer aos gestores a criação de painéis personalizados com os indicadores-chave de desempenho (KPIs) mais relevantes para suas necessidades, tais como: Tempo médio de resposta a emergências, Disponibilidade da frota, Custos operacionais e Número de ocorrências por tipo e região.

Este módulo ainda tem a responsabilidade da análise de dados históricos permitindo o estudo de tendências e padrões, auxiliando na identificação de áreas de risco, na otimização da alocação de recursos e na previsão de demandas futuras.

Oferece também a possibilidade de monitoramento em tempo real onde os gestores poderão monitorar a localização e o status de cada veículo da frota em tempo real, bem como os sinais vitais dos pacientes sendo transportados, garantindo uma resposta rápida e eficiente a incidentes críticos.

O sistema, através deste módulo, também poderá gerar alertas e notificações automáticas em caso de eventos críticos, como atrasos no tempo de resposta, falhas em equipamentos ou desvios de rota, permitindo uma intervenção imediata.

Finalmente a plataforma poderá gerar relatórios e dashboards personalizáveis para facilitar o acompanhamento do desempenho da frota e a comunicação de resultados para diferentes stakeholders

4.2 Persistência de dados

A camada de persistência é um componente importante do sistema, responsável por armazenar e gerenciar os dados gerados pelo aplicativo móvel, pelos dispositivos embarcados nas ambulâncias e pelas diversas fontes de informação integradas. A escolha do MongoDB como banco de dados se justifica pela sua capacidade de oferecer alto desempenho, escalabilidade e flexibilidade, requisitos essenciais para lidar com o grande volume e a variedade de dados gerados continuamente pelo sistema.

O MongoDB, um banco de dados NoSQL orientado a documentos, oferece diversas vantagens para este caso de uso, o armazenamento eficiente de Dados Não Estruturados e Semiestruturados: O MongoDB é capaz de armazenar de forma eficiente dados não estruturados e semiestruturados, como registros de sensores, dados biomédicos, eventos da frota e informações contextuais. Essa flexibilidade é crucial para lidar com a diversidade de dados gerados pelo sistema.

As operações de leitura e escrita de alta velocidade oferecidas pelo MongoDB são essenciais para manter a sincronização em tempo real do sistema e garantir a disponibilidade dos dados para os usuários e os algoritmos de análise.

De forma complementar, a escalabilidade horizontal, o que significa que novos nós podem ser adicionados ao cluster conforme o crescimento da demanda, garantindo que o sistema possa lidar com o aumento do volume de dados e do número de usuários sem comprometer o desempenho.

Este tipo de banco de dados possui um modelo de dados flexível, orientado a documentos do MongoDB permite que os dados sejam armazenados de forma adaptável, facilitando a evolução do sistema e a integração de novas fontes de dados.

5 Conclusão e trabalhos futuros

Este estudo explorou a aplicação da Internet das Coisas (IoT) na otimização do transporte de emergência, destacando a importância da integração de tecnologias para monitoramento, controle e gestão da frota de ambulâncias. A implementação de um sistema baseado em Web Service, Aplicativo Mobile, Banco de Dados e OBDII demonstrou potencial para aprimorar a eficiência operacional e a qualidade do atendimento emergencial, permitindo o rastreamento em tempo real dos veículos e a transmissão antecipada de dados críticos aos hospitais.

A adoção dessas tecnologias pode reduzir o tempo de resposta, melhorar a coordenação entre equipes de resgate e hospitais, proporcionando uma gestão mais inteligente dos recursos disponíveis. No entanto, desafios como infraestrutura de comunicação, interoperabilidade entre sistemas e segurança de dados ainda precisam ser endereçados para garantir a confiabilidade e escalabilidade da solução.

Este trabalho está em andamento e possui requisitos práticos a serem implementados, como a finalização das funcionalidades listadas nos capítulos anteriores, a implementação de um dashboard como forma de apresentar as ocorrências em tempo real. Outro aspecto fundamental é o avanço na modelagem das camadas relacionadas a emergências médicas para o gêmeo digital urbano, responsável por representar digitalmente, em tempo real, os eventos relacionados à mobilidade urbana e ao atendimento pré-hospitalar.

Adicionalmente a inclusão de funcionalidades no aplicativo mobile que permitam alterar os cenários e características de emergências.

Dessa forma, espera-se que o projeto evolua para um hub central de informações, além de poder ser utilizado por outros microsserviços, seguindo a filosofia de cidade inteligente, contribuindo significativamente para a modernização dos serviços de emergência urbana e garantindo uma resposta mais rápida e eficiente em situações críticas.

A longo prazo, a proposta visa consolidar-se para o gerenciamento inteligente de emergências, integrando múltiplas fontes de dados e fornecendo insights estratégicos que possam contribuir para a eficiência, agilidade e qualidade dos atendimentos.

Referencias

1. Abdeen, Mohammad AR et al. A Novel Smart Ambulance System—Algorithm Design, Modeling, and Performance Analysis. *Ieee Access*, v. 10, p. 42656-42672, 2022.

2. Adelantado, Ferran, et al. "Understanding the limits of LoRaWAN." *IEEE Communications magazine* 55.9 (2017): 34-40.
3. Adhav, Hrutunjay Balasabeb et al. Reduction of Ambulance Response Time and Accident Detection using IP and CPM.
4. Ayesha, Amreen; Chakravarthi, Komalavalli. Smart Ambulances for IoT Based Accident Detection, Tracking and Response. *Journal of Computer Science*, v. 19, n. 6, p. 677-685, 2023.
5. Batty, Michael, et al. "Smart cities of the future." *The European Physical Journal Special Topics* 214 (2012): 481-518.
6. Chiroli, Daiane Maria De Genaro et al. Integrating resilience and sustainability: A systematic analysis of resilient cities using ISO 37123. *International journal of disaster risk reduction*, v. 96, p. 103960, 2023.
7. Dalela, Pankaj Kumar, Smriti Sachdev, and Vipin Tyagi. "LoRaWAN network capacity for practical network planning in India." 2019 *URSI Asia-Pacific Radio Science Conference (AP-RASC)*. IEEE, 2019.
8. Fei, T. A. O. et al. makeTwin: A reference architecture for digital twin software platform. *Chinese Journal of Aeronautics*, v. 37, n. 1, p. 1-18, 2024.
9. Kristiningrum, E.; Kusumo, H. Indicators of smart city using SNI ISO 37122: 2019. In: *IOP Conference Series: Materials Science and Engineering*. IOP Publishing, 2021. p. 012013.
10. Madakam, Somayya, Ramya Ramaswamy, and Siddharth Tripathi. "Internet of Things (IoT): A literature review." *Journal of Computer and Communications* 3.5 (2015): 164-173.
11. Valdez, Larry et al. Design and development of an IoT-based smart ambulance system with patient monitoring. In: 2022 *IEEE International Power and Renewable Energy Conference (IPRECON)*. IEEE, 2022. p. 1-4.
12. Anagnostopoulos, Theodoros; Ntalianis, Klimis; Tsapatsoulis, Nicolas. IoT-enabled ambulances assisting citizens' well-being after earthquake disasters in smart cities. In: 2019 *International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. IEEE, 2019. p. 922-928.
13. Adibi, Sasan et al. Enhancing healthcare through sensor-enabled digital twins in smart environments: A comprehensive analysis. *Sensors*, v. 24, n. 9, p. 2793, 2024.



4º WORKSHOP SOBRE INDUSTRIA 4.0 Y HCI

Challenges and Strategies for Industry 4.0 Adoption in the Brazilian Wholesale Sector

Felipe Chuang¹[0000-0002-6699-8382], Maria Amelia Eliseo¹[0000-0003-0913-3259],
Daniela Vieira Cunha¹[0000-0001-6478-7138]
¹ Mackenzie Presbyterian University, São Paulo SP 01239-001, Brazil
felipe.chuang@mackenzista.com.br
{mariaamelia.eliseo, daniela.cunha}@mackenzie.br

Abstract. This study examines the challenges and strategies of small and medium-sized enterprises (SMEs) in the Brazilian wholesale sector for the adoption of Industry 4.0. A quantitative survey was conducted with SMEs in the southern and southeastern regions of Brazil, employing structured data collection and descriptive statistical analysis. The findings indicate that most SMEs remain in the early stages of digital transformation, with limited investments in innovation and challenges related to infrastructure, cultural adaptation, and financial constraints. The study underscores the need for financial incentives, workforce training, and the adoption of enabling technologies to enhance competitiveness and accelerate the digitalization process. Recommendations include targeted government policies, strategic partnerships, and increased access to technological resources to foster the integration of Industry 4.0 practices.

Keywords: Industry 4.0, Digital Maturity, SMEs, Digital Transformation, Technological Adoption.

1 Introduction

Industry 4.0 represents a transformative shift in the productive sector by integrating innovative technologies such as the Internet of Things (IoT), Big Data, Artificial Intelligence (AI), Cloud Computing, and Cybersecurity. Furthermore, it redefines how we work, communicate, and interact. These changes drive new approaches to management, production, and consumption [1], [2]. Process automation and optimization enhance efficiency, reduce costs, and facilitate real-time, data-driven decision-making [3].

While large enterprises have advanced in implementing Industry 4.0, small and medium-sized enterprises (SMEs) still face significant challenges, including limited funding, scarce investment in research and technological development, inadequate infrastructure, lack of familiarity with innovation, and, most notably, cultural resistance to adopting new technologies [4]. In Brazil, digitalization within this sector remains incipient, highlighting the need to identify barriers and establish strategies to accelerate this transformation.

To progress in this process, SMEs require an organizational culture that embraces innovation, employee training, investments in enabling technologies, and well-defined digital transformation strategies [5]. Additionally, adequate technological infrastructure, efficient data management, and access to financial resources are critical factors in overcoming these challenges. Thus, digital transformation is not an isolated goal but rather a continuous journey toward digital maturity for SMEs.

In this context, the primary objective of this study is to evaluate the adoption and digital maturity level of Brazilian small and medium-sized enterprises (SMEs) within the framework of Industry 4.0, focusing on companies located in the southern and southeastern regions of Brazil. The specific objectives are: (1) to analyze the factors influencing the adoption of these practices and (2) to propose recommendations to overcome the identified barriers. To achieve these objectives, a quantitative survey was conducted using a structured questionnaire, addressing aspects such as process digitalization, adopted technologies, and organizational culture. The findings provide insights into the adoption challenges and offer strategic recommendations to accelerate digital transformation in the sector.

This article is structured as follows: Section 2 presents the theoretical background on Industry 4.0 and digital maturity; Section 3 describes the adopted methodology; Section 4 discusses the obtained results; and finally, Section 5 presents the final considerations and suggestions for future research.

2 Theoretical framework

Industry 4.0 is an automation approach that leverages advanced technologies such as the Internet of Things (IoT), cloud computing, big data, artificial intelligence (AI), and robotics. These technologies enable factories to become more efficient, agile, and adaptable while also creating new business opportunities and service models [1]. The concept of Industry 4.0 goes beyond simple digitalization, introducing a new production and management paradigm based on interoperability and intensive data usage [1], [2], [6]. Industry 4.0 is not limited to the production stage but extends across the entire value chain, integrating manufacturing and distribution processes with after-sales services, including customer satisfaction, maintenance, and product or service loyalty [4].

Although its implementation presents challenges such as cultural resistance and the need for new investments, Industry 4.0 also offers significant opportunities for innovation and process optimization. Companies that have adopted Industry 4.0 solutions have reported up to a 30% reduction in production time, as well as significant cost savings and increased operational efficiency [7]. This demonstrates that adopting these technologies can serve as a key competitive advantage in the global industrial landscape.

However, the Brazilian industrial sector remains behind other countries in this transformation [4]. According to Brasil [8], fewer than 2% of Brazilian companies are fully immersed in Industry 4.0. Moreover, approximately 43% of companies do not recognize digital technologies as business drivers, which hinders industrial modernization. Nevertheless, estimates suggest that if Brazilian industries fully embraced digital transformation, the sector could generate an additional R\$ 73 billion per year [8].

It is also important to highlight the role of small businesses in the national economy. In 2020, they accounted for 27% of the Gross Domestic Product (GDP), and in 2021, they were responsible for approximately 4 million new businesses in Brazil [9]. Thus, Industry 4.0 should not be seen merely as a technological innovation but as an essential strategy for fostering growth and competitiveness among Brazilian companies, regardless of their size.

3 Methodology

This study is characterized as descriptive and exploratory research, employing a mixed approach (quantitative and qualitative). To achieve the proposed objectives, both quantitative methods, through a structured questionnaire, and qualitative methods, involving the analysis of guiding documents on digital innovation practices, were combined. This mixed approach allows for quantifying the degree of adoption of Industry 4.0 technologies in companies while exploring the perceptions of managers and employees regarding digital maturity and the challenges involved. Thus, it is possible to obtain a comprehensive view of the situation, integrating numerical data and perceptions about the context of digitization in small and medium-sized enterprises (SMEs) within the wholesale sector.

Data collection was carried out through an online questionnaire. The responses were analyzed using statistical techniques to identify patterns and trends in the adoption of Industry 4.0 practices. A descriptive analysis of the qualitative data was conducted to understand the motivations, barriers, and expectations of SMEs regarding digital transformation. Ethical research principles were followed, with all participants being informed about the research objectives, the voluntary nature of their participation, and the commitment to confidentiality of the information.

The research involved the participation of eight companies, selected through a convenience sampling method: five located in Rio de Janeiro, two in São Paulo, and one in Rio Grande do Sul. Among the companies

assessed, five are classified as small businesses, 62.5% of which have around 20 to 40 employees, while the remaining 37.5% employ more than 50 individuals. These companies operate in the wholesale sector, focused on the commercialization of various items, including home decorations, household goods, school supplies, and plush toys.

The questionnaire (Table 1) was designed based on the Brasil [10] and Oliveira [11]. These guides are tools that assist SMEs in adopting innovative practices and evaluating their digital maturity stage, offering specific questions that support the technological and strategic adaptation of businesses. The first document is focused on digital transformation and is structured into sections that help SMEs implement digital technologies and innovative practices. The second document provides a maturity model adapted to the reality of Brazilian SMEs, with a focus on practices applicable to industries across various sectors, including a case study in the animal feed industry.

Table 1. Digital maturity questionnaire for small and medium-sized enterprises in the wholesale sector (based on the Brasil [10] and Oliveira [11].

Question	Answers
1. How would you classify the investments in technological innovation over the past three years?	• None • Low • Moderate • High • Frequent • Nonexistent
2. At what stage do you believe your company will be, from a strategic and cultural perspective, in the next 5 years?	• Initial • Partial • Advanced • Complete • Nonexistent
3. How advanced is the digitization of your company's products (Internet of Things connectivity, digital monitoring, control, optimization, and automation)?	• Initial • Partial • Advanced • Fully Digitized • Nonexistent
4. How would you assess the preparedness of your company's people (resistance to change, openness to innovation, and skill development) for Industry 4.0?	• Initial • Partial • Advanced • Complete
5. Is there insecurity regarding the maintenance of jobs, or a sense of inadequacy to the new times?	• Yes • No • Partially
6. Does the company employ up-to-date technologies in its processes compared to market references?	• Yes • No • Daily • Weekly • Monthly • Semi-annually • Annually • Rarely • Never
7. How frequently does the company gather information and track market changes related to competitor actions and customer attitudes?	• Lack of financial resources • Lack of technical knowledge • Internal resistance to change • Inadequate infrastructure
8. What are the main challenges your company faces in adopting Industry 4.0 practices?	

9. Impact on competitiveness: Has the adoption of digital technologies impacted your company's competitiveness?
- Positively
 - Negatively
 - No impact

The questionnaire (Table 1) was structured into blocks of questions, each directed at a specific aspect of the companies' digital maturity:

- Questions 1 and 4: Aim to assess the level of investment in technological innovation. This information is relevant to understand the company's commitment to digital transformation and its willingness to allocate resources to Industry 4.0 technologies.
- Questions 2 and 5: Explore the company's strategic and cultural perspective regarding Industry 4.0. These items evaluate the companies' future vision and the degree of integration of digital technologies into organizational culture—key factors for the success of digitization.
- Questions 3 and 6: Investigate the degree of alignment of companies with the market in terms of technology and information. Both questions address the company's ability to remain competitive through innovation and market intelligence.
- Questions 7, 8, and 9: Analyze the relationship between the challenges faced by companies in adopting Industry 4.0 practices and the impact of these practices on competitiveness. Thus, the questions aim to connect the barriers to digital transformation with the expected outcomes in terms of business performance.

Each question or group of questions was designed to collect relevant and specific information about the companies' digital maturity, enabling a detailed analysis of the data obtained. The questions were constructed to be clear and direct, facilitating respondents' understanding and aiding in their responses.

3.1 Research limitations

Although more than 20 companies were invited to participate in the study, only eight agreed to take part. Consequently, the research was conducted with a limited number of companies, which may not fully capture the diversity and complexity of small and medium-sized enterprises (SMEs) within the context of Industry 4.0 in the South and Southeast regions of Brazil. Additionally, the participating companies belong to a specific market niche—the wholesale sector—dealing with a variety of products. This small sample size may limit the generalizability of the findings to other companies or industries. Furthermore, the responses may be subject to biases, as companies might have answered in a way they perceived as more favorable or may have underestimated or overestimated their practices and perceptions.

4 Results and discussion

This section presents and analyzes the results of the research on the digital maturity of SMEs within the context of Industry 4.0. The analysis provides an overview of the participating companies' profiles and the collected data, addressing key aspects such as investment in technological innovation, strategic and cultural positioning, digitalization level, and workforce preparedness.

To ensure the privacy of participating companies, the results are presented in an aggregated and anonymized manner, safeguarding their identities. The data were systematically organized and visually represented through graphs, facilitating clear and comparative analysis. Each graph is accompanied by an interpretation, outlining the companies' current digital maturity levels and highlighting the challenges and opportunities that arise in this context.

Analyzing the results of the applied questionnaire (Table 1), it is evident that 37.5% of companies classified their investment in technological innovation as "Low," while "Moderate" and "High" each accounted for 25%. Additionally, 12.5% of the companies reported no investment in innovation, and none of the participants indicated frequent investments in technology (Fig. 1). These findings suggest that, despite a willingness to

invest, most companies allocate insufficient resources to keep pace with Industry 4.0. This highlights the need for financial incentives to support innovation.

1. How would you classify the investments in technological innovation over the past three years?
8 respuestas

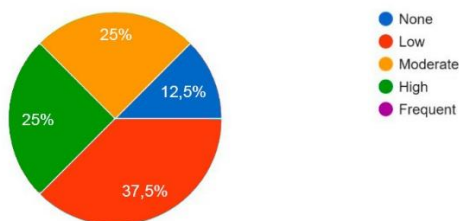


Fig. 1. Investments in technological innovation.

The Brazilian federal government has announced investments to help small businesses transition to digital operations, reduce costs, and improve efficiency [12]. However, for the effective adoption of Industry 4.0, companies must enhance their perception of the value of these investments. This can be achieved through training programs and specialized consultancy services that provide concrete data on the positive impact of digital transformation on business performance [13]. Awareness initiatives, such as workshops and case studies, can be effective in demonstrating the long-term benefits of technological innovation and convincing decision-makers of its strategic importance [14].

Regarding strategic and cultural perspectives for the next five years, 37.5% of companies anticipate remaining in the "Initial" or "Partial" stages, while only 25% have a more optimistic outlook and aim to reach the "Advanced" stage (Fig. 2). Many businesses are still in the early phases of digital adaptation, meaning they are just beginning to incorporate digital technologies into their processes. Concerns about change persist, leading them to rely on traditional methods with limited data integration and low automation [15]. Therefore, there is a lack of clear strategies for digital transformation, along with challenges related to infrastructure, limited financial resources, and insufficient awareness of technology's potential within the company. To address these barriers, it is important to invest in training programs for managers and employees, ensuring that digital strategies are effectively aligned with Industry 4.0 practices [16].

According to the responses, all companies in the sample have some level of digitalization, with 37.5% classified as "Advanced," while 25% are at the "Initial" and "Partial" stages, and only 12.5% have achieved full digitalization (Fig. 3). These findings indicate that, although most companies have implemented some degree of automation, few have reached complete digital transformation.

2. At what stage do you believe your company will be, from a strategic and cultural perspective, in the next 5 years?
8 respuestas

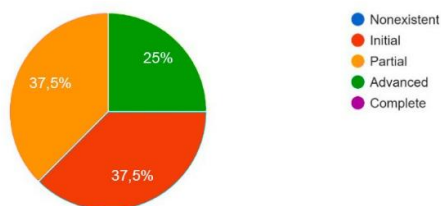


Fig. 2. Strategic and cultural stage of Industry 4.0.

According to Cintra et al. [5], product digitalization is one of the core pillars of Industry 4.0. For SMEs to achieve a high level of digital maturity, it is essential to invest in enabling technologies such as cloud computing, which eliminates the need for on-premises equipment, offering sustainability benefits such as energy savings, reduced physical space requirements, and lower infrastructure costs. Additionally, IoT (Internet of Things) enhances operational efficiency by introducing automation techniques and enabling a more flexible production system, while digital monitoring improves security and reduces operational costs.

3. How advanced is the digitization of your company's products (Internet of Things connectivity, digital monitoring, control, optimization, and automation)?
8 respuestas

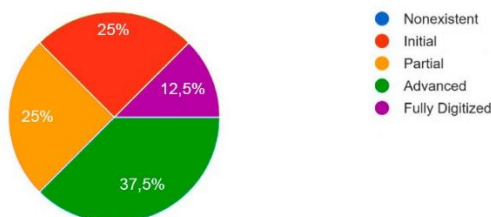


Fig. 3. Level of product digitalization.

Regarding workforce readiness for Industry 4.0, 25% of respondents rated their level of preparation as "Nonexistent," another 25% as "Initial," and 50% as "Partial" (Fig. 4). No company in the sample classified its readiness as "Advanced" or "Complete." These findings suggest that most companies are still in the early stages of cultural adaptation and workforce training for digital transformation.

The analysis highlights a gap in employee preparedness for adopting Industry 4.0 practices. This scenario reflects a cultural and strategic challenge that hinders the full integration of digital technologies within organizations. Organizational culture and employee willingness to embrace technological change are critical factors for the success of digital transformation [17]. Furthermore, Souza et al. [18] emphasize that continuous training and development programs are essential for overcoming resistance to innovation and equipping teams to operate in digital environments.

4. How would you assess the preparedness of your company's people (resistance to change, openness to innovation, and skill development) for Industry 4.0?
8 respuestas

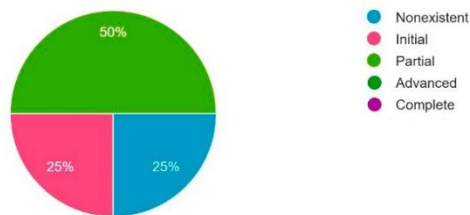


Fig. 4. Level of readiness for Industry 4.0.

The lack of preparedness identified reinforces the need for structured initiatives, such as specialized training programs, innovation workshops, and investments in both technical and soft skills. These efforts can enhance employee confidence, accelerate adaptation to technological changes, reduce internal resistance, and promote competitiveness in the Industry 4.0 landscape.

Regarding feelings of insecurity and inadequacy in adapting to modern technological changes, 25% of participants reported feeling uncertain about their preparedness for innovation, change, and skill development. In contrast, 37.5% stated they did not feel insecure, while another 37.5% reported partial insecurity (Fig. 5). These findings highlight an ambiguous scenario in which some companies do not feel fully prepared for the challenges and transformations of Industry 4.0.

5. Is there insecurity regarding the maintenance of jobs, or a sense of inadequacy to the new times?

8 respuestas

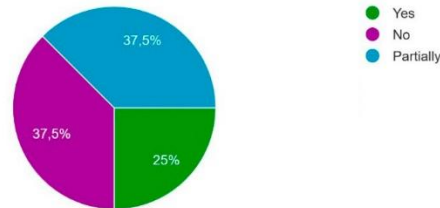


Fig. 5. Culture of adaptation to new technologie.

This result reflects the impact of digital transformation on labor market dynamics, which often raises concerns about whether current skills align with emerging demands. Recent studies indicate that Industry 4.0 introduces disruptive technologies such as automation and artificial intelligence, which may replace certain human roles while simultaneously creating opportunities for new high-skilled jobs [19].

To address these insecurities, training and reskilling programs are essential to better equip employees for future demands. According to Souza [6], integrating public and private policies focused on developing technological competencies and fostering an innovation-driven culture is crucial to reducing feelings of inadequacy and building confidence in the digital transition process. Additionally, organizational change management, with an emphasis on clear communication and psychological support, can help alleviate concerns about job obsolescence and promote a smoother adaptation to technological advancements.

The questionnaire results indicate that 75% of companies reported employing updated technologies in their processes, while 25% do not keep up with market standards (Fig. 6). These findings suggest a positive trend toward Industry 4.0, which relies on advanced technologies to enhance productivity and competitiveness.

6. Does the company employ up-to-date technologies in its processes compared to market references?

8 respuestas

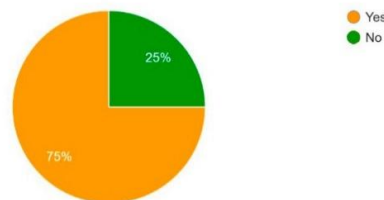


Fig. 6. Adoption of updated technologies in business processes.

The adoption of updated technologies, along with adaptability to change, keeps small businesses competitive in the market. Implementing Industry 4.0-enabling technologies, such as cloud computing and the Internet of Things (IoT), contributes to operational efficiency, enhances innovation capacity and business resilience, and

underscores that companies that have not yet embraced these practices are at a competitive disadvantage. Therefore, it is essential for such companies to consider strategic measures to bridge this gap [20].

Based on the questionnaire results, the frequency of data collection and market monitoring varies among companies. Only 25% conduct these activities daily, while the other intervals are evenly distributed at 12.5% for weekly, monthly, semiannual, and annual monitoring. Lastly, 25% of companies rarely engage in this practice (Fig. 7). This disparity highlights the lack of standardization in market intelligence practices, which may compromise business competitiveness.

According to Chaffey and Smith [21], continuous market analysis is a fundamental pillar for agile and data-driven strategies in the digital era. Moreover, the use of analytical tools and systematic processes can enhance companies' ability to anticipate trends and respond swiftly to market changes.

7. How frequently does the company gather information and track market changes related to competitor actions and customer attitudes?
8 respuestas

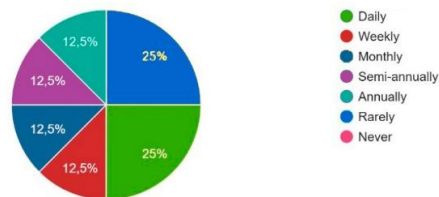


Fig. 7. Frequency of information collection and market monitoring.

Regarding the challenges faced by companies in adopting Industry 4.0 practices, 7.5% of companies indicated that internal resistance to change is the greatest challenge to adopting these practices, while 12.5% pointed to inadequate infrastructure (Fig. 8). These results highlight that the main barrier is related to cultural and behavioral issues, reflecting the difficulty companies face in promoting an innovative mindset and preparing their employees for digital transformation.

8. What are the main challenges your company faces in adopting Industry 4.0 practices?
8 respuestas

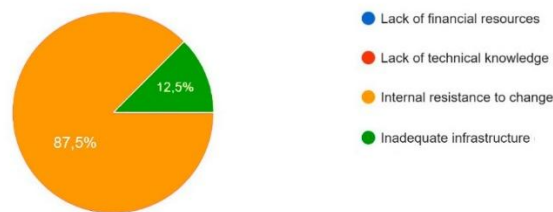


Fig. 8. Main challenges for the adoption of Industry 4.0 practices.

Resistance to change has been an impediment to digital transformation [15], [16] and [22]. Organizational change often encounters barriers related to human behavior, such as fear of the unknown, lack of clear communication, and insufficient involvement of leaders. These barriers can be mitigated through strategies such as effective communication, training, and engaging employees in the change process [23].

Furthermore, the inadequate infrastructure mentioned by 12.5% of companies points to structural challenges, such as the lack of integrated technological systems or connectivity, which are considered essential for implementing Industry 4.0 technologies like IoT and automation. Overcoming this challenge can be facilitated,

for instance, by government incentives, public-private partnerships, and strategic investments that enable access to enabling technologies [12] and [19].

Regarding the perception of whether the adoption of digital technologies has impacted the company's competitiveness, 50% of companies identify a positive impact from this adoption, while 37.5% indicate a negative impact, and 12.5% perceive that this adoption has not had a significant impact (Fig. 9). These results show that, although digital transformation is recognized as a lever for competitiveness, increased productivity, and resource optimization, its implementation presents some challenges.

9. Impact on competitiveness: Has the adoption of digital technologies impacted your company's competitiveness?
8 respuestas

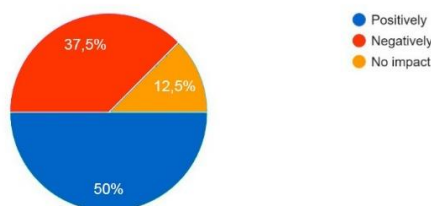


Fig. 9. Impact of digital technology adoption on business competitiveness.

This suggests that digital transformation also involves risks and challenges, such as high implementation costs, corporate conservatism, lack of internal competencies, and unfamiliarity with the use of technologies. Other risks can be highlighted, such as cybersecurity vulnerabilities, which involve exposure to new types of cyber threats, increasing the need for robust security measures; concerns about data privacy, as the storage of large volumes of data may raise questions about privacy and the protection of personal information; and technological integration challenges, which can become complex and costly, requiring significant planning and resources [24]. Moreover, the adoption of new technologies can lead to cultural changes within the organization, which may be difficult to predict and manage.

The positive impact observed in half of the companies aligns with the literature that highlights the benefits of digital transformation for increasing operational efficiency, innovation, and market differentiation. According to Chaffey and Smith [21], the adoption of digital technologies can enhance the customer experience, reduce operational costs, and promote greater agility and assertiveness in decision-making. Furthermore, digitalization facilitates access to new markets and increases the ability to adapt to changes in the competitive environment.

Based on the responses obtained from the questionnaire, despite the willingness of the companies in the sample to invest, the majority allocate insufficient resources to keep pace with Industry 4.0, highlighting the need for financial incentives to support innovation. Many companies are still in the early stages of digital adaptation, using traditional methods with limited data integration and low automation. There is a lack of clear strategies for digital transformation, along with challenges related to infrastructure, limited financial resources, and low awareness of technology's potential.

The disparity in technology adoption, with 75% of companies using updated technologies and 25% not keeping up with market standards, suggests a need for standardization and sharing of best practices. Companies that have yet to adopt advanced technologies could benefit from partnerships and collaborations with more advanced firms.

Companies perceive challenges in adopting Industry 4.0 practices, such as internal resistance to change (87.5% of companies indicated that internal resistance is the biggest challenge) and inadequate infrastructure (12.5% of companies pointed to inadequate infrastructure as a significant challenge). These challenges reflect

cultural and behavioral difficulties, hindering the promotion of an innovative mindset and the preparation of employees for digital transformation.

The data collected indicate that all companies reporting that the adoption of digital technologies negatively impacts their competitiveness are in the initial stage of digital transformation. These companies have little or no automation, their employees are unprepared (all respondents indicated that workforce readiness and openness to innovation are nonexistent or initial), and they rarely monitor market changes related to competitor actions and customer attitudes. One of these companies commented that 'more difficult than implementing new production processes is convincing managers of the need to reinvent themselves to stay in the market; more traditional companies are mostly more resistant to change, and this reflects in other sectors.'

Conversely, companies that believe they will achieve an advanced level from a strategic and cultural perspective in the next five years do not exhibit insecurity regarding job retention or feelings of inadequacy in relation to modern times. They are attentive to market changes and customer attitudes, although they also point to internal resistance to change

The results emphasize the importance of structured initiatives, such as specialized training programs, innovation workshops, and investments in both technical and interpersonal skills, to increase employee confidence, accelerate adaptation to technological changes, reduce internal resistance, and promote competitiveness in the Industry 4.0 landscape.

5 Final considerations

This study highlights the importance of digital transformation for the competitiveness of small and medium-sized enterprises (SMEs) in the context of Industry 4.0. The analysis enabled the identification of the digital profile of the participating SMEs and the assessment of key factors influencing their ability to innovate and adopt emerging technologies. The results indicate that, despite gradual progress in investment in innovation and digitalization, many companies still face strategic, cultural, and technological infrastructure challenges.

The study revealed that although investment in technological innovation is increasing, it remains insufficient for most SMEs to reach a competitive level of digitalization. From a strategic and cultural perspective, many of these companies are still in the early stages of digital transformation, highlighting the need for more comprehensive support, including workforce training and the alignment of digital strategies. Regarding product digitalization, the findings indicate that most of the analyzed companies have implemented some degree of automation, yet few have achieved full digital integration.

In light of these challenges, it is suggested that financial incentive policies, training programs, and greater access to enabling technologies be made more readily available to SMEs seeking digital transformation. Additionally, government incentives, strategic partnerships, and targeted investments in technology adoption could accelerate this process. These initiatives can support Brazilian SMEs in embracing Industry 4.0, enhancing their competitiveness and resilience.

For future research, it is recommended that new variables related to the digital transformation of SMEs be explored. While this study focused on the wholesale sector in the Southeast and South regions of Brazil, future research could broaden the scope to include SMEs from other economic sectors, such as industry and services, as well as companies located in less developed regions of the country. Such an approach would enable a more in-depth analysis of regional and sectoral specificities, contributing to a more comprehensive understanding of the challenges and opportunities associated with Industry 4.0 adoption.

Acknowledgments. This work was supported by MackPesquisa (Mackenzie Research and Innovation Fund [Fundo Mackenzie de Pesquisa e Inovação]), Project number 221006.

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

References

1. Schwab, K.: The Fourth Industrial Revolution. World Economic Forum (2016).
2. Eliseo, M. A., Silveira, I. F., Martins, V. F., Amato, C. A. de la H., Vieira, D., Carnevalli Junior, L., Cristiano, G. B., Chuang, F., Tanikawa, F.: The maturity of Brazilian companies on the adoption of industry 4.0 practices. *International Journal of User-System Interaction* 15(3), 51–68 (2022). <https://rochi.utcluj.ro/ijusi/articles/IJUSI-15-3-Eliseo.pdf>.
3. Milhossi, J. F., de Queiroga, A. P. G., Rodrigues, L. C.: Indústria 4.0 e a transformação digital. *Observatório de la Economía Latinoamericana* 21(12), 25566–25575 (2023). <https://doi.org/10.55905/oelv21n12-115>.
4. Passos, L. H. S.: A Indústria 4.0: fundamentos e principais impactos na economia brasileira. *Revista de Administração e Negócios da Amazônia* 12(2), 53 (2020). <https://doi.org/10.55905/rana.v12i2.53>.
5. Cintra, L. P., Braga, R. C., Rapini, M. S., Santos, U. P. dos, Barbosa, A. C. Q.: Indústria 4.0 e transformação digital: uma discussão conceitual, sob perspectiva neoschumpeteriana, que inclui políticas de CT&I e catch up. *Universidade Federal de Minas Gerais* (2019). DOI: 10.5752/P.1984-6606.2019v19n54p114-132.
6. Souza, E. M. de M.: Desafios da indústria 4.0 no contexto brasileiro: uma revisão de literatura. *ÍANDÉ: Ciências e Humanidades* 5(1), 44–57 (2021). DOI: 10.36942/iande.v5i1.128.
7. FIRJAN, Federação das Indústrias do Estado do Rio de Janeiro: Indústria 4.0 no Brasil: oportunidades, perspectivas e desafios (2029), <https://www.firjan.com.br/publicacoes/publicacoes-de-inovacao/industria-4-0-no-brasil-oportunidades-perspectivas-e-desafios.htm#pubAlign>.
8. Brasil, Ministério do Desenvolvimento, Indústria, Comércio e Serviços: Agenda brasileira para a indústria 4.0 (2018). <https://www.gov.br/mdic/pt-br/assuntos/noticias/mdic/mdic-e-abdi-lancam-agenda-brasileira-para-a-industria-4-0-no-forum-economico-mundial>.
9. Sebrae: Micro e pequenas empresas geram 27% do PIB do Brasil (2021). <https://sebrae.com.br/sites/PortalSebrae/ufs/mt/noticias/micro-e-pequenas-empresas-geram-27-do-pib-do-brasil,ad0fc70646467410VgnVCM2000003c74010aRCRD#:~:text=Bras%C3%ADlia%20%2D%20Os%20pequenos%20neg%C3%B3cios%20respondem,presidente%20do%20Sebrae%2C%20Luiz%20Barretto>.
10. Brasil, Confederação Nacional da Indústria: Guia de Referência para atuação em Gestão da Inovação com foco em Transformação Digital (2023). <https://www.portaldaindustria.com.br/publicacoes/2023/1/guia-de-referencia-para-atuacao-em-gestao-da-inovacao-com-foco-em-transformacao-digital/>.
11. Oliveira, G. A.: Modelo de maturidade para a indústria 4.0 para PME's brasileiras: um estudo de caso em uma indústria de ração animal. *Universidade Tecnológica Federal do Paraná* (2018). <http://repositorio.utfpr.edu.br/jspui/handle/1/4067>.
12. Agência Sebrae de Notícias, ASN: Governo federal anuncia R\$ 186,6 bilhões de investimentos na transformação digital da indústria. 11/09/2024. <https://agenciasebrae.com.br/economia-e-politica/governo-federal-anuncia-r-1866-bilhoes-de-investimentos-na-transformacao-digital-da-industria/#:~:text=Ser%C3%A3o%20R%24%207%20bilh%C3%B5es%20por,diretamente%20%C3%A0%20chamada%20ind%C3%BAstria%204.0>.
13. Saabye, H., Kristensen, T.B., Wæhrens, B.V.: Developing a learning-to-learn capability: insights on conditions for Industry 4.0 adoption. In: *International Journal of Operations & Production Management*, pp. 25–53. Emerald Publishing, (2022). <https://doi.org/10.1108/IJOPM-07-2021-0428>.
14. van Dun, D.H., Kumar, M.: Social enablers of Industry 4.0 technology adoption: transformational leadership and emotional intelligence. In: *International Journal of Operations & Production Management*, pp. 152–182. Emerald Publishing, (2023). <https://doi.org/10.1108/IJOPM-06-2022-0370>.
15. Gunji, B., Gondim, G., Leal, G., Sampaio, G., Eliseo, M.: Processo de desenvolvimento do Off-Line: um software de otimização de filas em praças de alimentação. In: *Anais do XXI Congresso Latino-Americano de Software Livre e Tecnologias Abertas*, pp. 212–221. SBC, Porto Alegre (2024). <https://doi.org/10.5753/latinoware.2024.245730>.
16. Mon, A., Del Giorgio, H. R.: Analysis of Industry 4.0 Products in Small and Medium Enterprises. In: *Procedia Computer Science*, pp. 914–923. Elsevier, (2022). <https://doi.org/10.1016/j.procs.2022.01.289>.
17. Andrade, M. E. G.; Macêdo, E. R.; Araujo, A. B.; Pereira, C. A.; Laurentino, G. K. S. C.; Lima, C. R. P.: Os impactos da cultura organizacional nos processos de trabalho: perspectivas diante da Revolução 4.0. *Revista Sociedade em Debate, Conselho de Ensino e Extensão, Faculdade Três Marias*, 4(2), 99–110 (2022). <https://sociedadeemdebate.com.br/index.php/sd/article/view/48/35>.
18. Souza, R. R. L.; Souza, T. J. S.; Mendonça, T. R.; Roberto, J. C. A.; Pinto Júnior, J. R. L.: A importância da capacitação e requalificação contínua do administrador para o alcance da valorização no mercado de trabalho atual frente à indústria 4.0. *Revista Científica Multidisciplinar Núcleo do Conhecimento*, 7(10), 35–45, (2022). <http://doi.org/10.32749/nucleodoconhecimento.com.br/administracao/capitacao-e-requalificacao>.
19. Centro de Gestão e Estudos Estratégicos – CGEE: Indústria 4.0: Arcabouço normativo para a implementação da Indústria 4.0 no Brasil. Centro de Gestão e Estudos Estratégicos, Brasília, DF, 29(214), 2021. https://www.cgee.org.br/documents/10195/11009594/cgee_sdt29_ind40_arcab.pdf/6cca2435-264b-47b0-9b57-d281e58c1b4b?version=1.1.

20. Zhou, K., Liu, T., Zhou, L.: Industry 4.0: Towards future industrial opportunities and challenges. In: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, pp. 2147-2152 (2015). <http://doi.org/10.1109/FSKD.2015.7382284>.
21. Chaffey, D., Smith, P. R.: Digital marketing excellence: planning, optimizing and integrating online marketing. Routledge, Oxfordshire (2022).
22. Maia, C., Santos, I. C., Sávio, M. A. C., Kubo, E. K. M.: Fatores da cultura organizacional que condicionam ou limitam o processo de inovação. Revista Capital Científico – Eletrônica (RCCe), 12(3), 2014. <http://doi.org/10.5935/2177-4153.20140027>.
23. Kotter, J. P.: Leading Change. Harvard Business Review Press, Boston (1996).
24. Bisultanova, A. A.: Threats and Risks of Digital Transformation. In: Modern Trends in Governance and Sustainable Development of Socio-economic Systems: From Regional Development to Global Economic Growth, pp. 482–489. European Proceedings of Multidisciplinary Sciences, London (2024). <http://doi.org/10.15405/epms.2024.09.54>.

Detección del Grado de Madurez en Aguacate Hass Usando Procesamiento Digital de Imágenes

Carlos Augusto Meneses Escobar¹, Julio Cesar Chavarro Porras¹,
Cesar Augusto Jaramillo Acevedo¹, Julio Cesar López Betancur¹

¹ Universidad Tecnológica de Pereira, código postal 660003, Pereira, Colombia.
{cmeneses, jchavar, swokosky, jucelo}@utp.edu.co

Resumen. El aguacate Hass es un fruto de interés en exportaciones de Colombia. Las condiciones de calidad y maduración óptima son fundamentales para garantizar un impacto positivo medioambiental y en la cadena de valor. Este trabajo presenta un modelo computacional basado en técnicas de procesamiento digital de imágenes para la medición del grado de madurez de cosecha, con base en las características físicas del aguacate Hass, a partir de repositorios de imágenes, que permiten una mejor clasificación y selección adecuada de los frutos. El procesamiento de imágenes contribuye significativamente a proporcionar un método preciso, objetivo y consistente para evaluar la madurez, reducir el error humano y respaldar el proceso general de control de calidad.

Palabras Clave: procesamiento digital de imágenes, grado de maduración, modelo computacional.

1 Introducción

El aguacate Hass es un fruto de origen centroamericano, representa más del 85% de los aguacates cultivados en el mundo. Es muy apetecido por su sabor y aportes nutricionales. Su importancia radica principalmente en su valor dietético y en gran medida al contenido nutricional y de ácidos grasos monoinsaturados, contenidos de minerales, vitaminas y otros fitoquímicos que contribuyen a la salud del ser humano [1].

El cultivo de Hass es nuevo en Colombia [2]. Desde el año 2009 se ha venido impulsando de manera decidida su producción, mediante las iniciativas privadas y con apoyo del sector público, transformándose en un pilar de desarrollo económico y social del país [3], [4]. En 2023 las exportaciones de aguacate Hass tuvieron un incremento del 17%, llegando a 114.535 ton y un incremento superior al 12% frente al año anterior superando los 200 millones de USD [5].

Colombia se ha convertido en el tercer exportador mundial de este fruto detrás de México y República Dominicana [6]. Esto plantea el desafío de producir en función de la calidad para el posicionamiento del producto a nivel internacional [7]. El panorama es promisorio para la industria aguacatera, sin embargo, la investigación y la adopción de tecnologías para el adecuado acompañamiento técnico del cultivo no crece en la misma proporción [8].

La calidad del aguacate es un reto importante [9]. Se pretende obtener frutas de alta calidad, para lograr mantener su buen estado por más tiempo, tal como lo exigen los mercados internacionales. Si el aguacate exportado presenta sobremaduración o residuos de productos no permitidos, es rechazado. Sin embargo, no se tiene suficiente información sobre indicadores de cosecha en Colombia, que permitan a los productores y exportadores protegerse adecuadamente para cumplir con los requerimientos del mercado [10]. Un estudio realizado [11] indica que los frutos de aguacate Hass presentan niveles considerables de rechazo por problemas relacionados con la calidad, que están directamente relacionados con prácticas inapropiadas de cosecha y postcosecha.

El aguacate es un fruto climatérico donde la madurez de consumo se produce después de la cosecha [12]. Sin embargo, la calidad y vida útil del fruto poscosecha están relacionadas con el periodo de tiempo desde su floración hasta lograr su madurez fisiológica, que coincide con el valor adecuado del contenido de materia seca [13]. Este periodo puede variar dependiendo de factores climáticos y atmosféricos, y de las prácticas agrícolas entre otros.

En un árbol se pueden formar cientos de frutos con diferencia de semanas, por lo no es fácil saber cuál es el momento exacto de formación, para determinar cuándo cosechar los frutos [14], [15]. Si un aguacate se cosecha inmaduro, no se ha formado lo suficiente y va a iniciar su maduración prematura [13], alterando su sabor y textura [16]. De otro lado, si permanece demasiado en el árbol, aumentan los problemas fitosanitarios y su vida útil se verá reducida.

La determinación precisa del estado de madurez del aguacate en cosecha, es un aspecto fundamental en la calidad de este. La decisión de cosechar un aguacate depende de la experticia del cosechador. En Colombia por ser un cultivo nuevo, muchos productores no tienen la suficiente experiencia y en su deseo de aumentar volúmenes de exportación, cosechan y agrupan frutos con diversos tiempos de formación.

Para medir el grado de madurez y calidad del fruto se emplean técnicas físicas y químicas. El contenido de materia seca (CMS), el contenido de aceite (CA) y el contenido de clorofila en la cáscara, se usan como índice de madurez [17]. Estas técnicas por ser técnicas destructivas, presentan desventajas pues requieren mucho tiempo para una preparación precisa de la muestra, además del error humano involucrado durante la selección aleatoria de muestras y del análisis manual. Las técnicas analíticas basadas en análisis de imágenes ofrecen ventajas en la detección (no invasiva) de los parámetros de madurez y de calidad del fruto del aguacate [18].

La falta de tecnologías automatizadas para determinar una cosecha adecuada del fruto es una problemática actual. Además, la determinación del estado de madurez óptimo del aguacate Hass en el momento de cosecha, se traduce en mayor calidad del fruto y ganancias económicas.

En la mayoría de soluciones propuestas a partir de imágenes digitales, no se especifican las técnicas utilizadas y en otras, no se indica el modelo computacional subyacente. No se conoce en el ámbito nacional, estudios soportados en modelos computacionales para determinar el nivel óptimo de madurez en cosecha de un aguacate Hass. Este trabajo presenta un modelo computacional que permite conocer el nivel de madurez de cosecha del aguacate Hass en el contexto colombiano, basado en técnicas de procesamiento digital de imágenes.

Optimizar y mejorar la precisión en la determinación del grado de madurez del aguacate Hass durante su cosecha, se hace necesario para mejorar los procesos de comercialización. Actualmente, los métodos tradicionales para evaluar la madurez pueden ser subjetivos y depender de la experiencia del productor, lo que puede resultar en variabilidad en la calidad del producto final.

El desarrollo de un modelo computacional basado en procesamiento digital de imágenes permitirá una evaluación objetiva y automatizada, reduciendo la incertidumbre en la toma de decisiones y garantizando que los frutos cosechados cumplan con los estándares de calidad requeridos para su comercialización. Además, esta tecnología podría agilizar el proceso de selección, minimizando pérdidas económicas y optimizando la eficiencia en la cadena de producción.

Este enfoque podrá contribuir a la implementación de técnicas de agricultura de precisión, facilitando el monitoreo y control de la madurez del aguacate de manera no invasiva. Con ello, se favorecerá la competitividad del sector agrícola al integrar herramientas tecnológicas avanzadas que permitan mejorar la calidad del producto y maximizar los rendimientos.

La determinación del estado óptimo de madurez del aguacate Hass en el momento de la cosecha se traduce en una mayor calidad del fruto y en ganancias económicas para el productor. Por ello, la aplicación de tecnologías automatizadas no invasivas, como el procesamiento digital de imágenes, permitiría al productor ofrecer un producto de alta calidad que garantice al consumidor un alimento con los nutrientes ideales.

Adicionalmente, la creación de un repositorio de imágenes de frutos, garantiza la estandarización en la evaluación de los frutos. Esto facilita la adopción de criterios uniformes en distintas regiones y mercados, mejorando la coherencia en las evaluaciones de calidad. Este repositorio puede ser utilizado como herramienta educativa para estudiantes y profesionales en áreas como visión por computador, IA aplicada y agroindustria, proporcionando un recurso práctico para el aprendizaje.

2 Estudios previos

Algunos estudios han permitido evaluar el comportamiento de los frutos durante el proceso de maduración poscosecha, basados en indicadores visuales de madurez como la luminosidad, color y matiz (tonalidad) de la cáscara y el tamaño, entre otros [19].

La clasificación de frutos para exportación se lleva a cabo mediante el aprendizaje automático o Machine Learning, y su objetivo es estudiar y construir algoritmos que pueden aprender para hacer predicciones sobre los datos recolectados. Algunos de los algoritmos más comunes son redes neuronales, regresión no lineal, agrupación jerárquica, árboles de decisión, lógica difusa, análisis discriminante lineal y procesamiento digital de imágenes [20].

El desarrollo de tecnologías automatizadas ha permitido la estimación de atributos de algunos productos hortícolas, basadas en características ópticas. Para los frutos de aguacate Hass se han implementado métodos no destructivos que incluyen técnicas basadas en imágenes espectrales [21] [22] y se han modelado los cambios de calidad durante el almacenamiento, mediante la implementación de una red neuronal artificial (ANN) y técnicas hiperespectrales [23].

Diversos estudios se enfocan en la evaluación de atributos de calidad específicos, pero no integran la evaluación cuantitativa de atributos de calidad externos e internos en un solo sistema. Actualmente, se han implementado diferentes estrategias con el fin de encontrar métodos no destructivos, rápidos, seguros y fáciles para la determinación del índice de madurez de cosecha de los frutos de aguacate Hass, basado en el procesamiento de imágenes.

Un estudio [24] presenta una metodología basada en procesamiento de imágenes mediante un sistema de visión computacional, para la evaluación de parámetros como color, firmeza y pérdida de masa relacionados con proceso de maduración de frutos de aguacate Hass. Igualmente, otra investigación desarrollada [22] propuso un algoritmo basado en el procesamiento digital de imágenes de frutos de aguacate Hass para la clasificación de acuerdo con su grado de madurez.

De otro lado, se ha realizado un estudio basado en el análisis de imágenes digitales mediante un sistema de visión computacional para mejorar el proceso de inspección de calidad de los frutos de aguacate Hass [25]. Otros estudios muestran la estimación del contenido de materia seca en frutos de aguacate Hass para la determinación de la madurez, mediante la implementación de un modelo de regresión de mínimos cuadrados parciales (PLS) con el uso espectroscopía infrarroja [26], y otro modelo usando espectroscopía [27].

Con el fin de abordar los desafíos asociados a la forma irregular, tamaño y colores variables de los frutos, se han realizado diversas investigaciones que comprenden adquisición de la información, su procesamiento, extracción de características y algo- ritmos de clasificación implementados en un solo sistema [25], [28].

Para la adquisición de datos, se utilizan técnicas no destructivas, que como se indicó anteriormente, es uno de los requisitos de mayor relevancia, y se adquieren en forma de imágenes, información espectral o valores discretos de tipo numérico o caracteres [29]. El procesamiento incluye una fase de preprocesamiento mediante la implementación de diferentes técnicas de estimación geométrica e intensidad, y una fase de segmentación para la extracción de la región de interés de una imagen basada en el brillo, color, escala de valores de grises, textura y bordes [30].

La extracción de características, se relaciona con descriptores de forma, mediante el uso parámetros geométricos que incluyen área, perímetro, varianza circular, diámetro, distancia radial, variancia elíptica, distancias máxima y mínima al centro de masas, entre otras [30], además, de las características de color que permite discriminar entre frutos, hojas, ramas y otros objetos de fondo, usando descriptores en formato RGB [31].

La clasificación se lleva a cabo mediante el aprendizaje automático [32] y su objetivo es estudiar y construir algoritmos que pueden aprender y hacer predicciones sobre los datos recolectados. Algunos de los algoritmos más comunes incluyen redes neuronales, regresión no lineal, agrupación jerárquica, árboles de decisión, lógica difusa, análisis discriminante lineal, entre otros [29].

Algunas técnicas de procesamiento digital de imágenes en formato RGB junto con redes neuronales han servido para monitorear los procesos de maduración de los aguacates Hass en poscosecha para ambientes controlados, corroborando su precisión cercana al 88% frente a indicadores de madurez, basados en contenido de materia seca y logrando predecir el tiempo en días para madurez de consumo en condiciones climáticas determinadas [33].

Diversos estudios se enfocan en la evaluación de atributos de calidad específicos, pero no integran la evaluación cuantitativa de atributos de calidad externos e internos en un solo sistema. Actualmente, se han implementado diferentes estrategias con el fin de encontrar métodos no destructivos, rápidos, seguros y fáciles para la determinación del índice de madurez de cosecha de los frutos de aguacate Hass, basado en el procesamiento de imágenes.

En la mayoría de soluciones propuestas a partir de imágenes digitales, no especifican las técnicas utilizadas y en otras, no se indica el modelo computacional subyacente. No se conoce en el ámbito nacional, estudios soportados en modelos computacionales para determinar el nivel óptimo de madurez en cosecha de un aguacate Hass.

3 Desarrollo metodológico

El objetivo del trabajo es desarrollar un modelo computacional basado en técnicas de procesamiento digital de imágenes que permita determinar con precisión el nivel adecuado de madurez de cosecha del aguacate Hass, para lo cual se siguió una metodología estructurada, en la que se determinaron las siguientes etapas:

Para la identificación de variables (como color, textura o brillo) para medir el grado de madurez del aguacate Hass basado en el análisis de imágenes digitales, se hace la especificación de la línea base de comparación determinada en análisis químico de algunas características para determinar el grado de madurez asociado a elementos como contenido de materia seca y clorofila. Una vez cosechados los frutos, se le hizo seguimiento al comportamiento del proceso de madurez a través de la medición de estas variables químicas.

Se tomaron muestras en el momento de cosecha de frutos en la zona Oeste de la cordillera Central y la zona Este de la cordillera Occidental del departamento de Risaralda (Colombia). Algunas muestras fueron sometidas a análisis químico en diferentes etapas de madurez para obtener la línea base de comparación. Además, se tomaron imágenes de frutos durante el proceso de maduración utilizando cámaras de dispositivos móviles convencionales para conformar una base de datos representativa.

En la fase de ajustes iniciales, a la imagen se le aplica un algoritmo de reducción de ruido donde se ajustan los pixels que son introducidos por errores del dispositivo que captura la imagen. Posteriormente, se hacen ajustes de contraste a la imagen. Los colores en una imagen se pueden ver alterados por el tiempo de obturación en la cámara o por el nivel de iluminación del objeto de interés al momento de capturar la imagen. Por lo tanto, es importante adecuar los colores a un rango válido para su procesamiento y análisis de los datos.

El procesamiento de la imagen se realiza en cuatro fases: Entrenamiento, Ajustes iniciales, Segmentación o preprocesamiento y Análisis.

En la fase de entrenamiento, se tomaron imágenes de aguacates para validar su grado de maduración correlacionando con variables físico-químicas. Esto con el fin de determinar los rangos válidos de valores en las variables del modelo, las cuales corresponden a coloración roja, verde y azul (extraídas del formato RGB), brillo, luminosidad.

Se determinaron algunas restricciones que se deben tener para garantizar la validez de la imagen a ser considerada. Entre estas, se tiene el nivel de iluminación de la imagen, el tamaño relativo del objeto de interés, la resolución y el tipo de imagen que se va a analizar, que en este caso es el formato RGB (ver Fig 1). Una imagen con mucha exposición a la luz o con bajo nivel de iluminación no permite identificar claramente el objeto de interés y además altera los valores de los datos asociados a este.

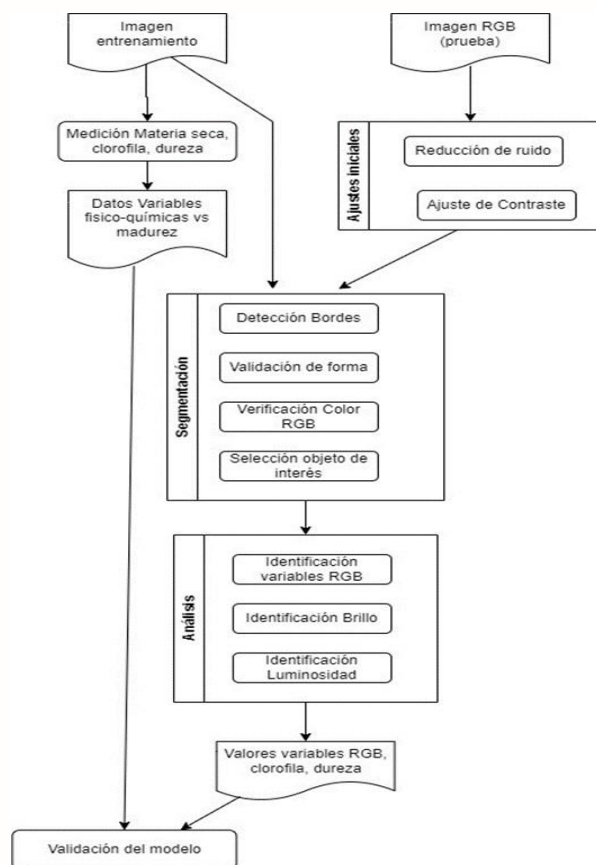


Fig. 1. Modelo para medir el grado de madurez (autor: propio).

Se realizó el preprocesamiento de imágenes, ajustando el color, eliminación de ruido y normalización para garantizar la calidad de los datos. Se valida la existencia del objeto de interés en la imagen, usando un algoritmo de detección de bordes y de validación de la forma cerrada que contiene la coloración adecuada del aguacate.

Se hace la extracción de características relevantes para identificar los valores de color y textura, para generar los datos de entrada a los algoritmos que entrenan el modelo. Se utilizaron algoritmos de aprendizaje con redes neuronales convolucionales para entrenar el modelo.

La validación y pruebas del modelo, se está realizando utilizando el conjunto de datos de prueba para medir su precisión comparando con la línea base dada por las variables materia seca y clorofila de acuerdo al grado de madurez, para ajustar parámetros según sea necesario. La implementación y optimización del modelo se está haciendo mediante un prototipo que usa dispositivos móviles, para optimizar su rendimiento en aplicaciones prácticas.

La fase de análisis permite determinar el grado de madurez del aguacate mediante la verificación de las variables de interés en el modelo, contrastadas con los rangos de los valores de las variables establecidas en el procesamiento de las imágenes de entrenamiento.

4 Experimentación

Las pruebas se realizaron en el laboratorio de Química de la Universidad Tecnológica de Pereira, con una temperatura promedio de 22 °C. Se fueron interviniendo día a día los aguacates de entrenamiento para calcular los valores físico-químicos con los que se contrasta el grado de madurez poscosecha.

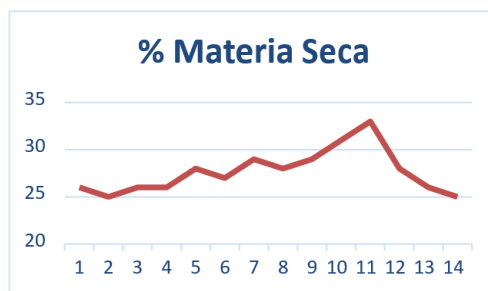


Fig. 2. Porcentaje de materia seca (autor: propio).

Se pudo determinar (ver Fig. 2) que el porcentaje de materia seca va aumentando en la medida que el aguacate va madurando, llegando a un valor máximo en su momento de consumo (en el día 11 aproximadamente). Posteriormente, este nivel baja de manera drástica por el ablandamiento de la pulpa al sobremadurarse. Tener un alto grado de materia seca al momento de cosecha es ideal para exportar, si se considera que en la medida que un aguacate madura, este porcentaje va aumentando.

El nivel de clorofila (medido en mg/g) y dureza decrecen en la medida que el fruto va madurando, (ver figuras 3 y 4). Igualmente, estas dos variables presentaron una caída importante contraria a la materia seca en el momento óptimo para consumo.

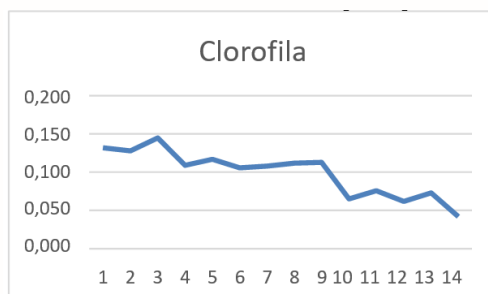


Fig. 3. Nivel de clorofila (mg/g) (autor: propio).

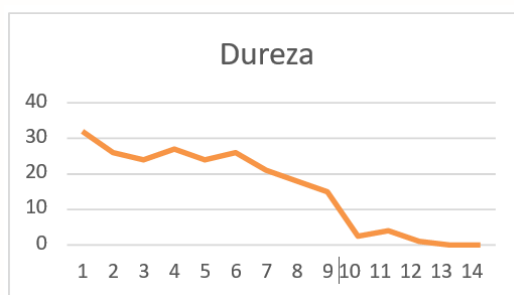


Fig. 4. Nivel de dureza (Kg/cm²) (autor: propio)

En la figura 5, se aprecian los cambios que ocurren en la coloración del aguacate Hass. Para cada color se mide en un rango de 0..255. En la medida que va madurando el fruto los valores de los colores convergen y la coloración verde permite contrastar de manera más clara el proceso de maduración frente a las variables físico-químicas anteriores descritas. A partir de esta imagen se puede determinar el rango adecuado de los colores que tiene un aguacate Hass para identificar el objeto de interés en la imagen procesada.

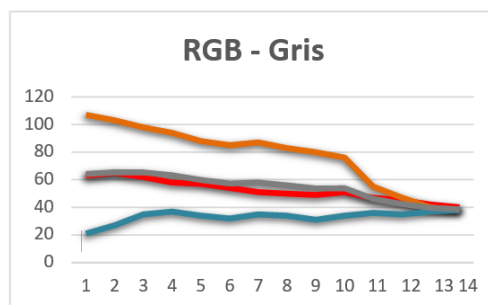


Fig. 5. Nivel de color RGB, gris (autor: propio).

Algunas pruebas preliminares del modelo propuesto usando redes neuronales convolucionales, han mostrado tasa de error cercana al con un 20%, en consideración con los estados inmaduro y maduro, debido a las similitudes de color que pueden presentar los frutos. Es importante destacar que el modelo aún está en proceso de desarrollo y se espera ajustar los resultados a las mediciones físico-químicas logradas en laboratorio.

5 Conclusiones y resultados

El modelo permite determinar el grado de madurez de un aguacate Hass y las pruebas contrastan en el mismo sentido con las variables físico-químicas.

Ajustando el proceso de entrenamiento, se podrá aplicar el modelo en otros entornos o regiones donde se pueda hacer seguimiento al proceso de maduración del aguacate Hass.

Uno de los resultados esperados con significativo impacto es la creación de un repositorio de imágenes que garantice la estandarización en la evaluación de los frutos, facilitando la adopción de criterios uniformes en distintos mercados y mejorando la coherencia en las evaluaciones de calidad.

Extensiones al modelo, deben considerar nuevas características físicas visibles que permitan medir la calidad del fruto, como el pedúnculo, la rugosidad, la forma o tamaño, entre otros.

Referencias

1. M. L. Dreher y A. J. Davenport, «Hass Avocado Composition and Potential Health Effects» *Crit. Rev. Food SCI Nutr*, vol. 53, nº 7, pp. 738-750, Jan 2013.
2. R. Uribe, «Cartama, casi dos décadas dedicadas al aguacate Hass,» *Semana*, 2 May 2018.
3. J. C. Diaz Vasquez, C. Ardila López y M. A. Guerra Aranguren , «Estudio de caso sobre la admisibilidad del aguacate Hass colombiano en el mercado estadounidense: oportunidades en el Este de Asia,» *Revista Mundo Asia Pacífico*, vol. 8, nº 14, 2019.
4. Minagricultura, «Memorias al congreso de la República 2018-2019,» 2017.
5. ANALDEX - Asociación Nacional de Comercio Exterior, «ANALDEX,» [En línea]. Available: [https://www.analdex.org/2024/03/18/informe-exportaciones-de-aguacate-hass-2023/#:~:text=Las%20exportaciones%20de%20aguacate%20Hass,con%20USD%20203%2C7%20millones](https://www.analdex.org/2024/03/18/informe-exportaciones-de-aguacate-hass-2023/#:~:text=Las%20exportaciones%20de%20aguacate%20Hass,con%20USD%20203%2C7%20millones.). [Último acceso: 8 jul 2024].

6. Analdex - Asociación Nacional de Comercio Exterior, «Analdex - Retos de producción, logísticos y climáticos, los principales dolores de cabeza para las exportaciones del aguacate Hass,» [En línea]. Available: [https://www.analdex.org/2024/05/29/retos-de- produccion-logisticos-y-climaticos-los-principales-dolores-de-cabeza-para-las-exportaciones-del-aguacate-hass/](https://www.analdex.org/2024/05/29/retos-de-produccion-logisticos-y-climaticos-los-principales-dolores-de-cabeza-para-las-exportaciones-del-aguacate-hass/). [Último acceso: 15 jul 2024].
7. J. G. Ramirez Gil, G. Franco y J. C. Henao Rojas , «Review of the concept of quality in Hass avocado and the pre-harvest and harvest factors that determine it under tropical conditions,» *Revista colombiana de ciencias hortícolas*, vol. 13, nº 3, 2019.
8. A. Hurtado Salazar y J. S. Arias Garcia, «Biología reproductiva y alternancia del aguacate Hass en el trópico andino de Caldas,» Manizales, 2023.
9. INFOBAE - Colombia potencia en producción de aguacate hass: qué le representa al país ser el tercer exportador del mundo después de México y República Dominicana, «INFOBAE,» [En línea]. Available: [https://www.infobae.com/colombia/2023/08/25/colombia-es-el-tercer-productor-mundial-de-aguacate-Hass/#:~:text=En%20Colombia%2C%20de%20acuerdo%20con,2021%20cerca%20de%2097.000%20toneladas](https://www.infobae.com/colombia/2023/08/25/colombia-es-el-tercer-productor-mundial-de-aguacate-Hass/#:~:text=En%20Colombia%2C%20de%20acuerdo%20con,2021%20cerca%20de%2097.000%20toneladas.). [Último acceso: 16 jul 2024].
10. P. Rodriguez , H. Juan Camilo, G. Correa y A. Aristizabal, «Identification of harvest maturity indicators for ‘Hass’ avocado adaptable to field conditions,» *Journal HortTechnology*, vol. 28, nº 6, pp. 815-821, 1 dic 2018.
11. J. G. Ramirez Gil, J. H. López y J. C. Henao Rojas , «Causes of Hass avocado fruit rejection in preharvest, harvest, and packinghouse: economic losses and associated variables,» *Journal Agronomy*, vol. 10, nº 1, p. 8, 19 12 2019.
12. V. Hershkovitz, H. Friedman, E. E. Goldschmidt, O. Feygenberg y E. Pesis, «Induction of ethylene in avocado fruit in response to chilling stress on tree,» *Journal Plant Physiology*, pp. 1855-1862, nov 2009.
13. M. d. M. Cerdas Araya, M. Montero Calderon y O. Somarribas Jones, «Verificación del contenido de materia seca como indicador de cosecha para aguacate (Persa americana) cultivar Hass en zona intermedia de producción de de los Santos Costa Rica,» *Agronomía Costarricense [online]*, vol. 38, nº 1, pp. 207-214, jan-jun 2014.
14. J. L. Rocha Arroyo, S. Salazar García y A. E. Bárcenas Ortega, «Determinación irreversible a la floración del aguacate ‘Hass’ en Michoacán,» *Revista Mexicana de Ciencias Agrícolas*, vol. 1, nº 4, 2010.
15. E. López Chaveli, «Efecto de la reducción de la floración con ácido giberélico sobre la producción del aguacate,» Valencia, 2020.
16. J. Gamble, R. Harker, S. Jaeger, A. White, C. Bava y M. Beresford, «The impact of dry matter, ripeness and internal defects on consumer perceptions of avocado quality and intentions to purchase,» *Postharvest Biology and Technology*, vol. 57, nº 1, pp. 35-43, Jul 2010.
17. P. J. Hofman, J. Bower y A. Woolf, «Harvesting, packing, postharvest technology, transport and processing. The avocado: botany, production and uses,» *CABI Digital Library*, pp. 489-540, 2013.
18. L. S. Magwaza y S. Tesfay, «A Review of Destructive and Non-destructive Methods for Determining Avocado Fruit Maturity,» *Food and Bioprocess Technology*, vol. 8, nº 10, pp. 1995-2011, 01 Oct 2015.
19. J. A. Herrera Gonzalez, S. Salazar Garcia, H. E. Martinez Florez y J. E. Ruiz Garcia, «Indicadores preliminares de madurez fisiológica y comportamiento poscosecha del fruto de aguacate Mendez,» *Fitotecnia Mexicana*, vol. 40, nº 1, pp. 55-63, 2017.
20. J. F. I. Opara y U. L. Nturambirwe, «Machine learning applications to non-destructive defect detection in horticultural products,» *Biosystems Engineering*, vol. 189, pp. 60-83, 1 Jan 2020.
21. S. Cubero, N. Aleixos, E. Moltó, J. Gómez Sanchis y J. Blasco , «Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables,» *Food and Bioprocess Technology*, vol. 4, nº 4 Springer, pp. 487-504, 24 May 2011.
22. E. R. Guerrero y G. Benavides, «Automated system for classifying Hass avocados based on image processing techniques,» *2014 IEEE Colombian Conference on Communications and Computing, COLCOM 2014 - Conference Proceedings*, 2014.
23. N. Maftoonzad, Y. Karimi, H. S. Ramaswamy y S. O. Prasher, «ARTIFICIAL NEURAL NETWORK MODELING OF HYPERSPECTRAL RADIOMETRIC DATA FOR QUALITY CHANGES ASSOCIATED WITH AVOCADOS DURING STORAGE,» *Journal of Food Processing and Preservation*, vol. 35, nº 4, pp. 432-446, Aug 2009.
24. I. Arzate Vásquez y et al., «Image Processing Applied to Classification of Avocado Variety Hass (Persea americana Mill.) During the Ripening Process,» *Food Bioprocessing Technology*, vol. 4, pp. 1307-1313, 2011.
25. H. Calvo , M. Moreno Armendáriz y S. Godoy Calderon, «A practical framework for automatic food products classification using computer vision and inductive characterization,» *Neurocomputing*, vol. 175, nº Part B, pp. 911-923, 29 Jan 2015.
26. B. B. Wedding, C. Wright, S. Grauf, R. D. White, B. Tilse y P. Gadek, «Effects of seasonal variability on FT-NIR prediction of dry matter content for whole Hass avocado fruit,» *Postharvest Biology and Technology*, vol. 75, pp. 9-16, an 2013.
27. J. A. Osuna García, P. Toivonen, S. Salazar Garcia, R. Goenaga y J. A. Herrera Gonzalez, «Modelo No Destructivo Para Determinar Madurez de,» *Memorias del V Congreso Latinoamericano del Aguacate*, vol. 04, pp. 420-428, 2017.
28. A. Gongal, S. Amatya, M. Karkee, Q. Zhang y K. Lewis, «Sensors and systems for fruit detection and localization: A review,» *Computers and Electronics in Agriculture - Elsevier*, vol. 116, pp. 8-19, 01 Aug 2015.

29. L. Nturambirwe y U. Opara, «Machine learning applications to non-destructive defect detection in horticultural products,» *Biosystems Engineering*, vol. 189 Academic Press, pp. 60-83, 01 Jan 2020.
30. K. Hameed , D. Chai y A. Rassau, «A comprehensive review of fruit and vegetable classification techniques,» *Image and Vision Computing*, vol. 80, pp. 24-44, Dec 2018.
31. Z. Kuś, J. Cymerski, J. Radziszewska y A. Nawrat, «Applying Colour Image-Based Indicator for Object Tracking,» *Studies in Systems, Decision and Control*, vol. 106 Springer International Publishing, pp. 23-33, 2018.
32. T. U. Rehman, M. S. Mahmud, Y. K. Chang y J. Jin, «Current and future applications of statistical machine learning algorithms for agricultural machine vision systems,» *Computers and Electronics in Agriculture*, vol. 156 Elseiver, pp. 585-605, 01 Jan 2019
33. C. A. Jaramillo, W. Choque, G. E. Guerrero y C. A. Meneses, «Hass avocado ripeness classification by mobile devices using digital image processing and ANN methods,» *International Journal of Food Engineering*, vol. 16, n° 12, p. 20190161, 28 Sep 2020

Emerging Technologies, HCI, and Agile Methodologies in Teacher Professional Development: A Study with 3D Printing and EduScrum

Eduar Rodolfo Becerra-Martínez¹ [0000-0002-5405-4624], Juan Manuel González-Calleros² [0000-0002-9661-3615], Erika Annabel Martínez Mirón² [0000-0002-4388-1624]

¹ Benemérita Universidad Autónoma de Puebla, Facultad Ciencias de la Electrónica, Puebla, México
bm223570612@alm.buap.mx

² Benemérita Universidad Autónoma de Puebla, Facultad Ciencias de la Computación, Puebla, México
jumagoca78@gmail.com, erika.mtz@correo.buap.mx

Abstract. This study presents the pilot implementation of a workshop that analyzes the impact of integrating eduScrum, 3D printing, TPACK, HCI principles, and User Experience (UX) in teacher professional development within the framework of integrative STEM (iSTEM), aligned with the principles of Industry 4.0 and Education 4.0. Through this experimental workshop, agile methodologies and emerging technologies were implemented to strengthen interdisciplinary teaching and analyze UX using 3D modeling software. The results show an increase in technological and pedagogical self-efficacy, with a notable emphasis on the Technological Pedagogical Engineering Knowledge (TPEK) dimension, suggesting that the combination of 3D printing and 3D modeling enhances STEM teaching. From an HCI and UX perspective, participants appreciated the software's accessibility, although some faced initial difficulties, highlighting the need for prior familiarization sessions. The integration of HCI principles optimized the UX and promoted a learning-centered design, facilitating technological appropriation. It is concluded that the combined use of eduScrum, 3D printing, and HCI principles supports teacher professional development in the context of Education 4.0.

Keywords: HCI, EduScrum, iSTEM, Education 4.0, Teacher Professional Development, Emerging Technologies.

1 Introduction

The Fourth Industrial Revolution has transformed how technology, automation, and digitalization converge across various sectors, including education [1]. In this context, Education 4.0 seeks to integrate emerging technologies such as artificial intelligence, augmented reality, and 3D printing into teaching and learning processes [2]. However, this integration should not be limited to incorporating digital or technological tools; it requires pedagogical models that effectively combine technology, methodology, and content.

One of the objectives of the 2030 Agenda for Sustainable Development (goal 4) is to ensure inclusive and quality education [3]. To achieve this, active methodologies such as flipped classrooms, project-based learning, and gamification have gained relevance, promoting autonomy and creativity in the classroom [4–6]. However, the lack of structured approaches creates a gap that hinders the effective integration of technology into educational processes while considering disciplinary and pedagogical content, forming an organized and structured set for educational processes [7].

Given this, the need for active methodologies to structure collaborative learning is evident. One such emerging and structured active methodology is eduScrum, which is based on the Scrum methodology [8]. Scrum is an agile methodology that organizes work into iterative cycles with defined roles, promoting self-efficacy and continuous improvement in teacher training [9].

Considering the integration of knowledge, the TPACK (Technological Pedagogical Content Knowledge) model provides a theoretical framework to understand the intersection of disciplinary, pedagogical, and technological knowledge [10], enabling better inclusion of emerging technologies in the classroom. In the case

of 3D printing, teachers must balance these knowledge domains (technology, pedagogy, and content) to design effective didactic strategies in STEM areas (Science, Technology, Engineering, and Mathematics).

Moreover, the interaction between users, 3D modeling technology, and the final product involves Human-Computer Interaction (HCI) and User Experience (UX), facilitating the design of accessible and effective educational resources [11]. The usability of 3D modeling software and ease of learning are critical factors in teachers' technological appropriation.

Based on the above, this study focuses on analyzing how 3D printing, TPACK, HCI, and eduScrum contribute to integrative STEM (iSTEM) education, which is the interdisciplinary approach to STEM [12]. To this end, an experimental workshop was designed, where 13 pre-service and in-service teachers worked in teams to develop didactic proposals based on 3D modeling. Key aspects evaluated included the ease of use of the software, self-efficacy in STEM, and collaborative work patterns.

Study Objectives

- Identify the impact of HCI UX principles on the appropriation of 3D modeling and printing technologies.
- Analyze how 3D printing enables the application of STEM knowledge in teaching.
- Evaluate the effectiveness of eduScrum in organizing learning based on emerging technologies.

This study aims to provide empirical evidence on the implementation of 3D printing and agile methodologies in teacher professional development for iSTEM, consolidating an educational model aligned with Education 4.0 and Industry 4.0.

2 Technological integration in education

In recent years, Education 4.0 has driven the integration of emerging technologies into teaching and learning processes, aligning with the needs of Industry 4.0. Among these technologies, 3D printing, HCI, and agile methodologies have gained relevance as innovative strategies to enhance STEM skill development and improve teacher and student self-efficacy.

The STEM approach has been widely recognized for its ability to develop critical thinking and problem-solving skills. The incorporation of 3D printing into this approach has enabled more interactive and tangible learning experiences, fostering interdisciplinarity and project-based learning [13]. Studies such as [14] have highlighted how additive manufacturing contributes to students' technical skill development and the appropriation of emerging technologies. Within the TPACK framework, literature has pointed out the importance of balancing disciplinary, pedagogical, and technological knowledge to design effective educational experiences for both STEM and iSTEM [13, 15]. In this sense, 3D printing not only facilitates the acquisition of technical knowledge but also allows teachers to strengthen their pedagogical knowledge by designing innovative educational materials [16].

Furthermore, HCI has been fundamental in optimizing digital tools in education. Research over the years has emphasized the importance of usability and user-centered design in adopting learning technologies. Studies [18] and [19] have demonstrated that implementing usability principles in learning experience (LX) improves UX and learning retention. Particularly in the context of 3D modeling software, user-centered design facilitates student interaction and participation in digital creation processes, enabling greater immersion through intuitive and accessible interfaces [20]. Recent research has analyzed how user interfaces (UI) should be efficient and tailored to users' needs to maximize the adoption of emerging technologies in education [21, 22].

Regarding agile methodologies, eduScrum has been identified as an innovative approach that promotes autonomy and collaborative work in education. Studies [23] have shown that this methodology structures learning into iterative cycles, providing clarity in student roles and objectives. Research by [9] indicates that eduScrum optimizes teamwork dynamics by fostering self-regulation and organization. Additionally, some studies have demonstrated that combining agile methodologies with emerging technologies can enhance personalized teaching and student motivation [24]. Similarly, [25] have analyzed how these methodologies

foster self-directed learning and interaction with teachers and peers, consolidating more dynamic and effective teaching processes.

Despite advances in integrating emerging technologies, HCI, and agile methodologies in education, there remains a gap in the literature regarding the convergence of these approaches within a structured framework to enhance iSTEM teaching. Various studies have emphasized the need for educational models that combine emerging technologies and agile methodologies with TPACK to strengthen active learning and improve teacher self-efficacy [15, 26]. Research by [24] and [2] highlights that this integration can facilitate the creation of more efficient and adaptive learning environments, while [25] emphasize the importance of designing strategies that combine agile methodologies and emerging technologies to foster creativity and critical thinking in students.

3 Methodology

This study followed a mixed-methods approach [27] to analyze the integration of eduScrum, iSTEM, TPACK, and HCI in an educational setting to enhance teacher professional development and improve their self-efficacy in integrative STEM (iSTEM). The methodology combined quantitative and qualitative techniques to gain a comprehensive understanding of the workshop's impact on teacher training and the appropriation of emerging technologies.

3.1 Study design

This study followed a sequential explanatory design, a mixed-methods approach that combines quantitative and qualitative data collection. The study was conducted in two phases:

- Quantitative Phase: This phase involved measuring changes in teacher self-efficacy in iSTEM using pretest and posttest assessments.
- Qualitative Phase: Following the quantitative analysis, a retrospective session was conducted to explore teacher perceptions, experiences, and reflections on integrating active methodologies and emerging technologies.

The study incorporated the TPACK framework to assess how teachers integrated technology into their pedagogical practices. The workshop structure balanced technological knowledge (TK), pedagogical knowledge (PK), and content knowledge (CK) to ensure meaningful learning experiences. Additionally, the study applied Human-Computer Interaction (HCI) and User Experience (UX) principles to evaluate the usability and accessibility of 3D modeling tools in education.

3.2 Participants

The workshop included 13 teachers, both pre-service and in-service, from diverse disciplines and regions affiliated with a pedagogical university in southern and southeastern in Mexico. Participation was voluntary, and a demographic profile was collected to understand the participants' backgrounds and experiences with technology integration in education.

Demographic breakdown:

- Gender distribution: 53.8% female, 46.2% male.
- Age range: 77% were between 20 and 30 years old, while 23% were over 30.
- Role in education: 76.9% were pre-service teachers, and 23.1% were in-service teachers.
- Teaching experience: 23% had more than 10 years of experience, 15.3% had between 2 and 10 years, and 61.7% had less than one year.
- Familiarity with STEM tools: Only one participant had prior experience with STEM tools.

This profile allowed researchers to assess the level of familiarity participants had with technology and their potential for adopting it in educational settings.

3.3 Workshop development

The workshop lasted four hours and was structured based on the agile eduScrum methodology, incorporating elements of Design Thinking to foster creative problem-solving in educational contexts using 3D printing.

The workshop was equipped with computers and two 3D printers for experimentation, in addition to a Kanban board that allowed for the visual management of workflow progress. The general phases of the workshop were:

- a. **Welcome and Introduction:** The "Promote" strategy from the INSPIRA model of "Tecnológico de Monterrey" [28] was used to create an engaging and motivating environment for participants. Initial exploratory questions on STEM, 3D printing, and 3D modeling were posed.
- b. **3D Modeling and Design:** Participants were introduced to the Tinkercad 3D modeling software, analyzing its user interface from an HCI and UX perspective. Usability, accessibility, and user-centered design aspects were highlighted, emphasizing the importance of intuitive interfaces for facilitating adoption by teachers.
- c. **Preparation for 3D Printing:** Discussions were held on the applications of 3D printing in Education 4.0 and Industry 4.0, linking them with active methodologies and showcasing successful cases in STEM education.
- d. **Designing Didactic Activities Using eduScrum:** Participants worked in teams, organized using eduScrum principles, and assumed key roles such as:
 - **Scrum Master:** Facilitated communication and removed obstacles.
 - **Product Owner:** Ensured that activities aligned with educational objectives, representing the end-users (students).
 - **Designer/Creator:** Focused on the 3D modeling of didactic prototypes.
 - **Technology Specialist:** Managed the operation of 3D printers.
 - **Group Evaluator:** Applied assessment instruments and guided the retrospective process.

Through the application of Design Thinking, each team followed these stages:

- **Empathize and Define:** Identification of educational problems and review of relevant literature.
 - **Ideate:** Brainstorming to generate didactic solutions.
 - **Prototype:** Creation of 3D models in Tinkercad.
 - **Evaluate and Refine:** Peer review of designs to optimize the models.
- e. **Reflection and Closing:** A retrospective session, based on the Scrum methodology, was conducted where participants reflected on the overall process, the knowledge acquired, and opportunities for improvement.

3.4 Data collection instruments

To assess the impact of the workshop on teacher self-efficacy in iSTEM, the TP-STEMK survey [29] was applied. This internationally validated instrument measures teacher self-efficacy in four key dimensions:

- **Technological Pedagogical Scientific Knowledge (TPSK):** Evaluates the use of scientific knowledge and processes in teaching with technology.
- **Technological Pedagogical Mathematical Knowledge (TPMK):** Measures the application of technology in mathematical instruction.
- **Technological Pedagogical Engineering Knowledge (TPEK):** Assesses the relationship between technology, disciplinary content, and engineering processes in education.
- **Integrative STEM (iSTEM):** Examines the interdisciplinary integration of STEM processes in the classroom.

Additionally, qualitative data were collected during the retrospective phase to analyze participants' perceptions of user experience (UX) with 3D modeling software and their views on usability and accessibility in implementing 3D printing in education.

This mixed-methods approach ensured a comprehensive evaluation of both quantitative improvements in self-efficacy and qualitative insights into teacher experiences and challenges with emerging technologies.

4 Results

Due to time constraints, the proposed didactic projects could not be fully materialized through 3D printing. However, participants demonstrated creativity by presenting sketches and 3D models that illustrated three specific didactic proposals, showcasing the educational potential of 3D technology across different disciplinary contexts. These proposals were:

- **Support System for Learning the Central Nervous System:** The team "Los Pedagógicos" designed an interactive didactic material to facilitate the learning of the central nervous system (CNS) through 3D printing. Their primary objective was to allow students to recognize the different parts of the nervous system using a layer-by-layer detachable model that reveals the human anatomy and, progressively, the internal structures of the CNS.
- **Enhancing Fine Motor Skills Through 3D Geometric Models:** The team "Los Naranjitas" developed an interactive didactic material aimed at improving fine motor skills in primary school students through an artistic and geometric approach. By using 3D-printed regular geometric shapes with central holes, the objective was to have children manipulate and fit the pieces together, thereby stimulating motor development and improving hand and finger precision.
- **Learning Decimal Numbers and Fractions Through 3D Models:** The team "Socialización" created an interactive didactic material called "pizza-pizza" to teach decimal numbers and their fractional equivalents. Using 3D-printed pieces that represent fractions of a unit, students could visualize and manipulate relationships between halves, quarters, eighths, and sixteenths, facilitating a tangible understanding of fractions.

These results are illustrated in Fig. 1.



Fig. 1. Didactic proposals generated by the teams.

In terms of measuring self-efficacy in iSTEM using the TP-STEMK survey, statistical analyses, including descriptive statistics, mean comparisons, correlations, and regression analysis, were performed using Python to determine the influence of each dimension on teacher self-efficacy.

4.1 Descriptive analysis

Means and standard deviations were calculated for each dimension of the TP-STEMK model (TPSK, TPMK, TPEK, and iSTEM) in pretest and posttest applications. The results indicate an overall increase in means across all evaluated dimensions. Notably, the TPEK dimension (Technological Pedagogical Engineering Knowledge) exhibited the most significant improvement, with a pretest mean of 3.85 (± 2.07) and a posttest mean of 4.00 (± 1.58). The variability in responses, measured through standard deviation, showed a tendency to decrease in the posttest, suggesting a more homogeneous perception of self-efficacy after the workshop. Figure 2 displays the comparison of means for each dimension, revealing an overall improvement in iSTEM self-efficacy after the intervention. Additionally, Figure 3 illustrates a general reduction in standard deviation for each dimension, indicating more consistent responses in the posttest and suggesting a decrease in variability in teachers' perceptions of self-efficacy in STEM education with the TPACK framework.

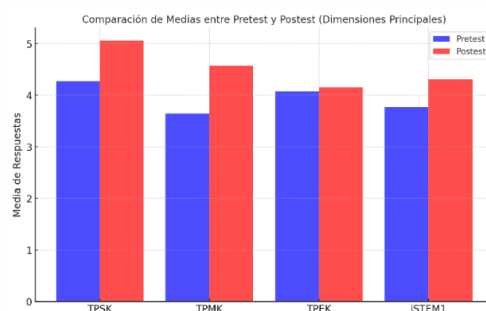


Fig. 2. Mean scores by dimension for pretest and posttest.

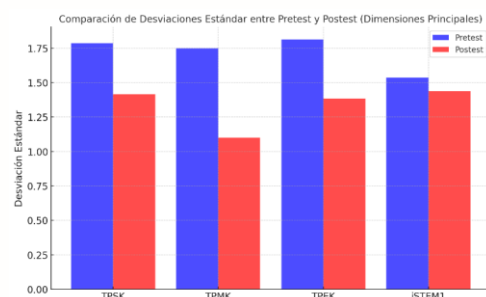


Fig. 3. Standard deviation by dimension for pretest and posttest.

4.2 Comparison of means between pretest and posttest

To evaluate the significance of changes in self-efficacy, statistical tests for mean comparisons were applied. Depending on the normality of the data (assessed using the Shapiro-Wilk test), a paired t-test was used for normally distributed data, while the Wilcoxon signed-rank test was applied for non-normally distributed data.

The results indicated statistically significant differences in most dimensions:

- **TPSK:** $t = -0.222$, $p = 0.829$
- **TPMK:** $t = -1.185$, $p = 0.266$
- **TPEK:** $t = 4.441$, $p = 0.002$
- **iSTEM:** $t = 2.964$, $p = 0.003$

These results demonstrate a significant increase in teacher self-efficacy across most dimensions following the pedagogical intervention. Figure 4 illustrates the overall change in self-efficacy from pretest to posttest.

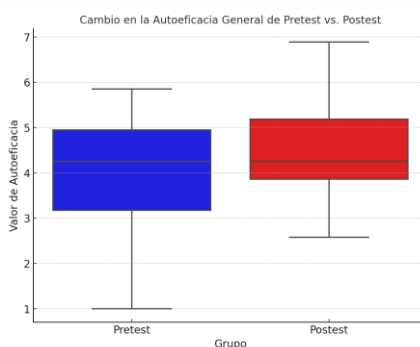


Fig. 4. Overall change in self-efficacy pretest vs. posttest.

4.3 Correlation analysis

To explore the relationships between the TP-STEMK dimensions and self-efficacy in iSTEM, Pearson and Spearman correlation coefficients were calculated. The results revealed that TPEK was the most strongly correlated dimension with iSTEM ($r = 0.87$, $p < 0.01$), suggesting that technological pedagogical engineering knowledge is a key predictor in the integration of STEM. TPSK and TPMK showed moderate correlations with iSTEM ($r = 0.54$ and $r = 0.43$, respectively).

The detailed correlation between each general dimension and iSTEM development is illustrated in Figure 5.

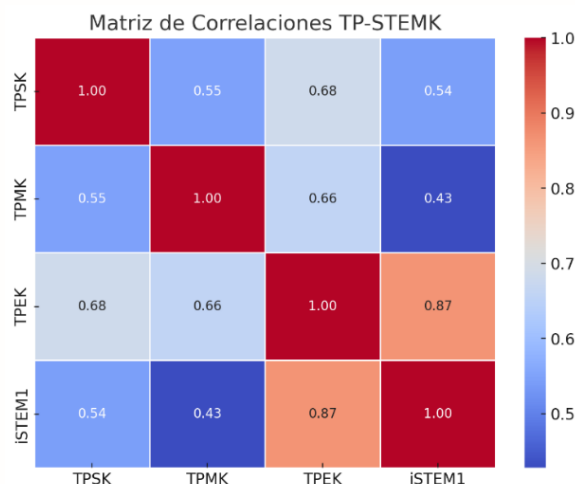


Fig. 5. Correlation matrix between TP-STEMK dimensions.

4.4 Regression analysis

To determine the predictive influence of each dimension on iSTEM, a multiple linear regression analysis was conducted. The adjusted model explained 78.7% of the variance in iSTEM ($R^2 = 0.787$, $p < 0.01$). The coefficient for TPEK was the only statistically significant predictor ($\beta = 1.10$, $p = 0.002$), indicating that this dimension has a direct effect on self-efficacy in iSTEM, while TPSK and TPMK did not show significant effects. Figure 6 presents scatter plots showing the linear relationship between TPEK, TPSK, TPMK, and iSTEM, while Figure 7 displays a radar chart highlighting the contribution of each item to self-efficacy in iSTEM.

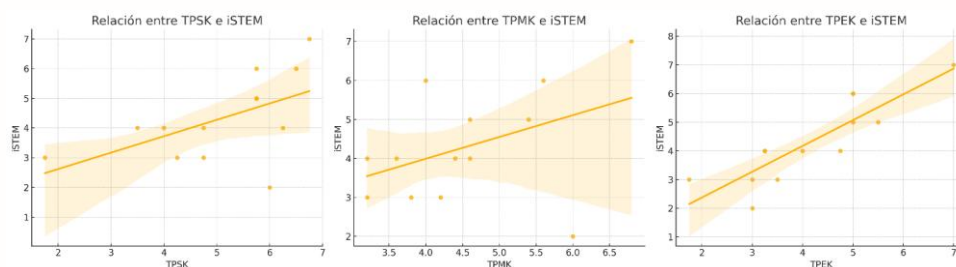


Fig. 6. Relationship between TPEK, TPSK, TPMK, and iSTEM.

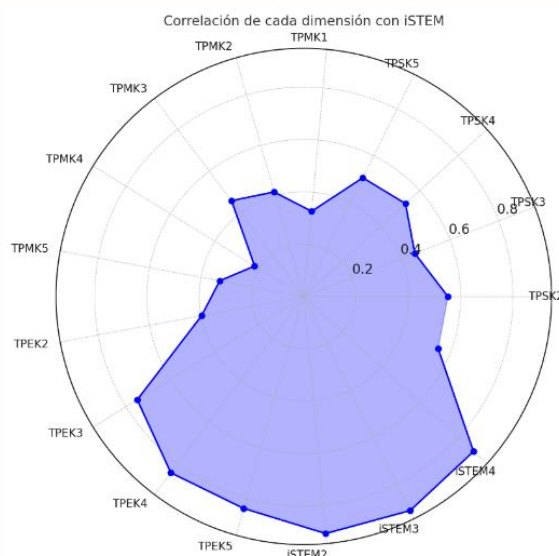


Fig. 7. Contribution of each item to self-efficacy in iSTEM.

4.5 Retrospective analysis

The retrospective analysis conducted at the end of the workshop identified strengths, areas for improvement, and proposals for optimizing future sessions. Participant feedback reflected the intersection of the STEM model, technology integration through TPACK, the user-centered approach from HCI and UX, and the structured learning process with eduScrum.

Participants highlighted that eduScrum structured collaborative work effectively, clarifying roles and optimizing learning management. The user experience (UX) in 3D modeling was positively received, although some found the learning curve challenging, suggesting the need for preparatory exercises. From an HCI perspective, the accessibility of the software was appreciated, but some teachers faced initial difficulties, emphasizing the need for onboarding strategies. The integration of iSTEM and TPACK enabled participants to apply technological knowledge (TK) to design educational models, enhancing the comprehension of abstract concepts. The combination of active methodologies and emerging technologies reinforced teacher self-efficacy in 3D printing-based instruction.

The main challenges identified included unstable internet connectivity, which limited access to online resources and slowed down 3D modeling, causing frustration among less tech-savvy teams. Additionally, the lack of prior training with the software led to delays and a steeper learning curve, impacting UX by reducing exploration time. From an HCI perspective, it became evident that strategies are needed to lower entry barriers to educational software, such as tutorial guides or intuitive onboarding to enhance accessibility and optimize participant experience.

Participants proposed several actions to improve the workshop experience in future editions:

- More practice with 3D modeling software: Participants suggested extending the time dedicated to exploring the software before starting the actual modeling process.
- Greater dissemination of the eduScrum methodology: Teachers believed this methodology could be applied in other educational contexts to promote structured and collaborative learning.
- Extended workshop duration: Some participants felt the time was insufficient to complete the design and evaluation cycle, proposing longer or multi-session formats.

Additionally, teachers suggested some adjustments to improve the workshop dynamics and optimize learning outcomes:

- Longer exploration phase before modeling: A more extensive initial phase was recommended to familiarize participants with digital tools before applying them in instructional design.
- Additional planning time: Before starting the modeling phase, participants suggested incorporating a research phase where they could explore examples of 3D printing applications in different educational disciplines.

Interestingly, no elements of the workshop were suggested for removal, indicating that the structure was well received and that the available time was used effectively.

5 Discussion

The results indicate that the integration of eduScrum and 3D printing positively impacted self-efficacy in TPACK for iSTEM. The TPEK dimension recorded the highest increase, aligning with previous studies that highlight the importance of technological knowledge in iSTEM [14]. These findings support the idea that project-based learning and the use of emerging technologies strengthen teachers' confidence in the application of digital tools.

However, the TPSK and TPMK dimensions showed less impact, as evidenced by the analysis of Figures 6 and 7, where greater dispersion is observed in their relationship with iSTEM development. This suggests the need for specific strategies to integrate technology into science and mathematics teaching. Previous research has documented that teachers require concrete examples to effectively connect technology with disciplines such as mathematics and science [12]. In future implementations, specific modules could be designed to guide the didactic application of 3D printing in these areas.

From the perspective of HCI and UX, participants valued the accessibility of the 3D modeling software, although some experienced initial difficulties. These findings are consistent with studies that emphasize the importance of usability in the adoption of educational technologies [19]. Implementing preliminary familiarization sessions with technological tools could optimize user experience and reduce learning barriers.

The eduScrum methodology facilitated the organization of collaborative work and project-based learning. The use of visual tools, such as the Kanban board, was key to managing the process and promoting teacher autonomy, validating its effectiveness in educational environments [9, 23]. However, participants suggested extending the duration of the workshop to deepen each phase.

As limitations, this study acknowledges the small sample size and the absence of a control group in this pilot test, indicating that future studies should expand the population and compare traditional methodologies with innovative approaches by using control and experimental groups to better evaluate their impact on STEM teaching.

6 Conclusions

This study demonstrates that the integration of eduScrum, 3D printing, iSTEM, TPACK, HCI, and UX in teacher professional development enhances technological and pedagogical self-efficacy, particularly in the TPEK dimension. The project-based methodology and the application of emerging technologies enabled teachers to develop practical skills in 3D modeling and printing, strengthening their ability to integrate these tools into their didactic strategies.

The findings highlight the importance of providing preliminary familiarization activities with 3D modeling software, as well as structuring workshops with extended durations to allow for progressive learning. Additionally, the evaluation of UX in the use of the software and the organization of work through eduScrum were key factors for the success of the intervention.

At the theoretical level, this study provides evidence of the relationship between self-efficacy in iSTEM and the implementation of agile methodologies, supporting previous literature on the effectiveness of eduScrum in Education 4.0. On a practical level, the results suggest that the combination of these strategies can be applied in other educational contexts to promote iSTEM teaching through digital design and fabrication.

For future directions, it is recommended to conduct studies with larger samples and experimental designs that allow for a more precise evaluation of the methodology's impact. Furthermore, the application of these strategies could be explored across different educational levels and disciplines, broadening their scope within educational innovation in the digital age.

References

- Schwab, K.: The Fourth Industrial Revolution. Crown (2017).
- de OLIVEIRA, L.C., Guerino, G.C., de OLIVEIRA, L.C., Pimentel, A.R.: Information and Communication Technologies in Education 4.0 Paradigm: a Systematic Mapping Study. *Informatics in Education*. 22, 71–98 (2023). <https://doi.org/10.15388/infedu.2023.03>.
- Gobierno de México: ¿Qué es la Agenda 2030 para el Desarrollo Sostenible?, <http://www.gob.mx/inafed/articulos/que-es-la-agenda-2030-para-el-desarrollo-sostenible>, last accessed 2024/03/14.
- Escrig-Tena, A.B., Segarra-Ciprés, M., García-Juan, B., Salvador-Gómez, A., Flor-Peris, M.L., Alexandra Badoiu, G.: Evaluación de una experiencia de flipped classroom. *TCE*. 33–70 (2023). <https://doi.org/10.51302/tce.2023.3399>.
- Hernández-Horta, I.A., Monroy-Reza, A., Jiménez-García, M.: Aprendizaje mediante Juegos basados en Principios de Gamificación en Instituciones de Educación Superior: Learning through Games based on Principles of Gamification in Higher Education Institutions. *Formación Universitaria*. 11, 31–40 (2018). <https://doi.org/10.4067/S0718-50062018000500031>.
- Secanell, I.L., Segura, F.R.: El arte como eje vertebrador de proyectos en educación infantil y primaria. Estudio de caso de colaboración escuela-universidad. *Rev. complut. educ.* 34, 389–400 (2023). <https://doi.org/10.5209/rced.79328>.
- Tejada, J., Thayer, T.: Diseño, implementación y evaluación de una intervención de formación en tecnología musical basada en TPACK y ABP en la formación inicial del profesorado de música de Educación Secundaria. *Revista Latinoamericana de Tecnología Educativa - RELATEC*. 18, (2019). <https://doi.org/10.17398/1695-288X.18.2.9>.
- Sutherland, J., Sutherland, J.J.: Scrum. Ariel, España (2018).
- Wijnands, W., Stolze, A.: Transforming Education with eduScrum. In: Parsons, D. and MacCallum, K. (eds.) *Agile and Lean Concepts for Teaching and Learning: Bringing Methodologies from Industry to the Classroom*. pp. 95–114. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-2751-3_5.
- Mishra, P., Koehler, M.J.: Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*. 108, 1017–1054 (2006). <https://doi.org/10.1111/j.1467-9620.2006.00684.x>.
- Lazar, J., Feng, J.H., Hochheiser, H.: *Research Methods in Human-Computer Interaction*. Morgan Kaufmann (2017).
- Chai, C.S.: Teacher Professional Development for Science, Technology, Engineering and Mathematics (STEM) Education: A Review from the Perspectives of Technological Pedagogical Content (TPACK). *Asia-Pacific Edu Res.* 28, 5–13 (2019). <https://doi.org/10.1007/s40299-018-0400-7>.
- Aguirre-Muñoz, Z., Pando, M.: Conceptualizing STEM teacher professional knowledge for teaching ELs: Initial impact of subject matter and disciplinary literacy PD on content knowledge and practice. *Bilingual Research Journal*. 44, 335–359 (2021). <https://doi.org/10.1080/15235882.2021.1970654>.
- Harron, J.R., Emert, R., Thomas, D.M., Campana, J.: Laying the Groundwork for STEAM: Scaling and Supporting 3D Design and Printing in Higher Education. *Front. Educ.* 6, (2022). <https://doi.org/10.3389/feduc.2021.763362>.
- Chaipidech, P., Kajonmanee, T., Chaipah, K., Panjaburee, P., Srisawasdi, N.: Implementation of an Andragogical Teacher Professional Development Training Program for Boosting TPACK in STEM Education: The Essential Role of a Personalized Learning System. *Educational Technology & Society*. 24, 220–239 (2021).
- Ford, S., Minshall, T.: 3D printing in education: a literature review. (2016).
- Kefalis, C., Skordoulis, C., Drigas, A.: The Role of 3D Printing in Science, Technology, Engineering, and Mathematics (S.T.E.M.) Education in General and Special Schools. | EBSCOhost, <https://openurl.ebsco.com/contentitem/doi:10.3991%2Fijoe.v20i12.48931?sid=ebsco:plink:crawler&id=ebsco:doi:10.3991%2Fijoe.v20i12.48931>, last accessed 2025/02/16. <https://doi.org/10.3991/ijoe.v20i12.48931>.
- Cooper, M., Colwell, C., Jelfs, A.: Embedding accessibility and usability: considerations for e-learning research and development projects. *Research in Learning Technology*. 15, (2007). <https://doi.org/10.3402/rlt.v15i3.10933>.
- Choi, G.W., Seo, J.: Accessibility, Usability, and Universal Design for Learning: Discussion of Three Key LX/UX Elements for Inclusive Learning Design. *TechTrends*. 68, 936–945 (2024). <https://doi.org/10.1007/s11528-024-00987-6>.

20. Nofal, E.: Studying User Experience in Architectural Education Using Unity Game Engine. In: Ibrahim, A., Mohamed, M.A.A., and Fekry, M. (eds.) *Man and Place*. pp. 33–43. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-49903-6_3.
21. Maldonado, T.V.P.: Diseño de Interfaces de Usuario Centrado en el Usuario: Mejores Prácticas y Tendencias. *Revista Académica y científica VICTEC*. 4, (2023).
22. Zheng, H., Xiao, D., Zhou, J.: Enhancing Perceived Value in Human-Computer Interaction: The Mediating Role of User Participation and the Moderating Role of Task Complexity. *International Journal of Human-Computer Interaction*. 1–10 (2024). <https://doi.org/10.1080/10447318.2024.2350839>.
23. Wijnands, W., Fritsch, K.: Experience the 'Flows' with eduScrum: About the Why, How, and What of the Empowerment of Students and Teachers. In: *Industry Practices, Processes and Techniques Adopted in Education*. pp. 219–243. Springer, Singapore (2022). https://doi.org/10.1007/978-981-19-3517-6_11.
24. Bala, M.Y., Damla, K.: A Review of Human-Computer Interaction Design Approaches towards Information Systems Development. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*. 12, 229–250 (2021). <https://doi.org/10.70594/brain/12.1/180>.
25. Li, K.C., Wong, B.T.: Agile-Blended Learning: An Innovative Technology-Mediated Approach to Distance Education. *SN COMPUT. SCI*. 6, 127 (2025). <https://doi.org/10.1007/s42979-024-03646-y>.
26. Chai, C.S., Rahmawati, Y., Jong, M.S.-Y.: Indonesian Science, Mathematics, and Engineering Preservice Teachers' Experiences in STEM-TPACK Design-Based Learning. *Sustainability*. 12, 9050 (2020). <https://doi.org/10.3390/su12219050>.
27. Creswell, J.W.: *Research design: qualitative, quantitative, and mixed methods approaches*. Sage Publications, Thousand Oaks, Calif (2009).
28. Tec. de Monterrey: MODELO INSPIRA, <https://sites.google.com/tec.mx/lineamientossectec/7/7-5-1-modelo-inspira>, last accessed 2025/02/23.
29. Chai, C.S., Jong, M., Yin, H., Chen, M., Zhou, W.: Validating and Modelling Teachers' Technological Pedagogical Content Knowledge for Integrative Science, Technology, Engineering and Mathematics Education. *Journal of Educational Technology & Society*. 22, 61–73 (2019).

Recommendations Guide for Adopting Industry 4.0 Practices for Small and Medium-Sized Brazilian Companies in the Food Sector

Gabriel Batista Cristiano¹[0000-0002-9539-1352], Fabio Tanikawa¹[0000-0003-0866-5755],
Huizilopoztli Luna-García²[0000-0001-5714-7482], Maria Amelia Eliseo¹[0000-0003-0913-3259]

¹Laboratório de Tecnologias Interativas (TecInt), Universidade Presbiteriana Mackenzie, São Paulo SP 01239-001, Brazil

²Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX), Universidad Autónoma de Zacatecas, Jardín Juárez 147, Centro Histórico, 98000, Zacatecas, México.
gabriel.cristiano@mackenzista.com.br, vinifaro.fabio@gmail.com,
hlugar@uaz.edu.mx, mariaamelia.eliseo@mackenzie.br

Abstract. This article aims to provide guidance on the adoption of processes and implementation of specific Information and Communication Technologies (ICT) products with the aim of supporting the technological transformation of small and medium-sized companies in the food sector within industry 4.0, thus improving their competitiveness in the Marketplace. Industry 4.0 studies were identified and evaluated using the systematic mapping method. To complement this analysis, interviews were carried out with companies. The results were used to create a recommendation guide for small and medium-sized companies, with the aim of guiding the adoption of Industry 4.0 technologies and processes.

Keywords: Industry 4.0, small and medium-sized enterprises, Brazil, recommendation guide, processes.

1 Introduction

In recent years, with technological advancements and the digitalization of products and services, the way business is conducted and how consumers engage with products has also changed. This transformation in production methods is known as the Fourth Industrial Revolution or Industry 4.0. According to [1], Industry 4.0 is defined by connectivity, enabling easy access to data. This is the primary distinction between the factories of the future—companies that have embraced Industry 4.0—and traditional legacy production-line factories that have yet to make the transition.

In Brazil, Industry 4.0 is still in a developmental phase and "remains significantly behind compared to other countries" [2]. According to [3], less than 2% of companies are immersed in Industry 4.0. Additionally, it is estimated that approximately 43% of companies do not recognize digital technologies as business drivers, despite the same study projecting that the industry could generate an annual value of BRL 73 billion if it adopted a full-scale digital transformation. When analyzing the Brazilian production landscape, small businesses play a vital role in the economy, contributing 27% to the Gross Domestic Product (GDP) in 2020 and accounting for approximately 4 million new businesses in 2021 [4]. It is also essential to promote the digital transformation of these enterprises to ensure they are not left behind by the innovations driven by Industry 4.0, making them even more relevant and competitive in the Brazilian market. However, small businesses must overcome barriers such as the high costs associated with adopting new technologies and the need to acquire the necessary knowledge to implement the innovations that enable Industry 4.0 ecosystems.

To identify and select best practices and recommendations for managing the adoption and implementation of specific information and communication technology (ICT) products that drive technological transformation for small and medium-sized enterprises (SMEs) within Industry 4.0, a systematic mapping method was conducted as shown on the [5]. This method was used to establish a theoretical foundation for developing a recommendation guide tailored to the reality of Brazilian SMEs. Based on the collected material, questions were formulated to conduct interviews with SMEs to understand their relationship with Industry 4.0 technologies and practices.

This paper is structured as follows: Section 2 presents the theoretical framework and related works. Section 3 presents the methodology of the work. Section 4 presents the results achieved. Section 5 demonstrates the creation of the guide. Section 6 shows the discussion, and Section 7 provides the final considerations and directions for future research.

2 Theoretical framework

This section discusses digital maturity, Industry 4.0 and its ecosystem, and innovation management.

2.1 Definition of Industry 4.0

Industry 4.0 emerged as a response to Asian market competition, representing a revolution in manufacturing by incorporating digital technologies and management concepts. By promoting communication and data sharing across all layers of the ecosystem, it prioritizes interoperability. Technologies such as the Internet of Things (IoT), Big Data, and Artificial Intelligence (AI) reshape processes and drive innovation. To remain competitive, digital transformation requires limited resources and the ability to overcome resistance to change. As suggested by Machado (2021), adopting Industry 4.0 necessitates a reassessment of conventional business models. [6] emphasizes that these changes are essential for the development of smart and efficient factories ([1],[2], [6],[7],[8]).

2.2 Related work

Among the recommendations considered in this study, the contribution of [9] stands out, providing a starting point through a literature review on the development of Industry 4.0 within small and medium-sized enterprises (SMEs) in Brazil. The authors report that Industry 4.0 development in Brazil is still in its early stages, primarily restricted to large corporations. However, by analyzing the country's production landscape, they highlight the critical role played by both small and large businesses. They argue that promoting digital transformation among SMEs is essential to keeping them competitive in the global market.

Additionally, the authors emphasize that overcoming barriers to SME adoption of Industry 4.0—such as high technology costs and a lack of awareness of its benefits—requires joint initiatives between public and private organizations, as well as collaboration with professionals in IT, management, and even Human-Computer Interaction (HCI). These efforts are necessary to prepare and guide SMEs through their transformation.

[7] exploratory-descriptive study investigated digital transformation awareness among small hospitality business managers in São Luís, Maranhão. The study aimed to identify related activities, technological adoption challenges, and future perspectives of these managers. The findings revealed that while managers have some prior knowledge of digital transformation and use common technologies, they still see full adoption as a future goal, despite recognizing its importance and benefits. The study concludes by highlighting the need to move past outdated concepts, as digital transformation is crucial for organizations' growth and market positioning.

[10] analyzes the impact of adopting Industry 4.0 technologies on the innovation of SME business models. The research includes multiple case studies, featuring interviews with experts and SME representatives. The results reveal several factors influencing SME adoption of Industry 4.0 technologies and their impact on business innovation. The study underscores the role of technological understanding in SME business model innovation and competitiveness. The findings provide valuable insights for managers, stressing the importance of technology for competitiveness and suggesting policies for technology adoption to foster SME development and market positioning.

This paper differentiates itself from related studies by proposing a structured set of recommendations for adopting Industry 4.0 practices and ICTs, specifically tailored to the realities of small and medium-sized

businesses in São Paulo's food sector. The goal is to assess their digital maturity level and guide them toward adopting beneficial practices that support their growth and competitiveness.

3 Methodology

This chapter describes the methodology adopted for this research, highlighting its qualitative nature by following the procedures of the systematic mapping process to gather the necessary literature and support the construction of the recommendation guide. It also emphasizes its quantitative aspect by collecting and analyzing data from interviews conducted with SMEs.

3.1 Development of the research instrument

To assess the relationship of SMEs with Industry 4.0 technologies and processes applied to their daily activities, an online questionnaire was developed to gain insights into their experiences, challenges, and practices related to adopting Industry 4.0 technologies. This approach aimed to correlate reports on the practices and technologies used by these companies with the recommendations obtained from the systematic mapping studies. The results of this systematic mapping are presented in [5].

The questionnaire was designed based on studies from [11] and the Brazilian Agency for Industrial Development [12], which provide guides with question repositories for SMEs to assist them in adopting Industry 4.0 practices. These questions are broad and applicable to companies of different sizes and maturity levels, encouraging entrepreneurs to analyze their businesses. The [11] study categorizes its questions into five groups: vision and direction, culture and engagement, structure and processes, resources, and infrastructure. Meanwhile, the [12] study organizes its questions into categories defined by the research team for better understanding: data-driven organization, customer value and connection, competitive foundations, and enabling technologies. Together, both studies contributed to a repository of 115 questions.

The selected categories and questions were filtered based on the SME context, particularly their resource limitations. Each question was analyzed for its relevance to SMEs, considering their industry, size, and available time. Ultimately, 21 questions were chosen and distributed into five new categories defined by the research team to cover the areas highlighted in the reference studies:

- Technological Infrastructure, Tools, and Applications: Comprising four questions, this category aims to understand SMEs' current relationship with essential Industry 4.0 technologies.
- Automation and Integration: Containing two questions, this category seeks to determine the level of data integration within business processes, evaluating automation and system integration capabilities.
- Management and Organizational Culture: With four questions, this set aims to assess whether company management adopts Industry 4.0 practices and whether there is a way to measure the application of these technologies.
- Strategy and Innovation: Including four questions, this section explores whether companies actively seek innovation, how they intend to leverage it, and how they promote their services.
- Market and Customers: The largest category, with seven questions, focuses on understanding the company's market presence, customer interactions, awareness of competitors, and use of technological tools to stay competitive.

These questions aim to deepen the analysis and provide a broader perspective on Industry 4.0 technology adoption and impact on SMEs. Beyond the theoretical recommendations from the systematic mapping studies, they contribute to constructing a recommendation guide that reflects the real relationship between SMEs and Industry 4.0 technologies.

3.2 Research participants

For the interviews, the research team established selection criteria for participating companies. They had to meet the following requirements:

- Employ between 10 and 99 workers to qualify as an SME, according to Sebrae's classification [4].
- Be located within or around the Higienópolis campus of the Mackenzie Presbyterian University.
- Operate in the food industry, to tailor recommendations to a specific business reality.

Interviews were conducted with 10 companies, selected through convenience sampling. The research team primarily interviewed business owners or employees responsible for managing the establishment's operations. Interviewees were assured that their company information would remain confidential, and participation was voluntary.

3.3 Development of the recommendation guide

The development of the Industry 4.0 best practices guide for SMEs was based on two sources of information: the systematic literature mapping conducted in the first phase of this research, with results published in [5], and the data collected from interviews with food sector companies, conducted in the second phase, which is the focus of this article.

The systematic mapping allowed for the identification and analysis of 19 relevant articles, providing a theoretical foundation for Industry 4.0 practices. These articles were examined to extract insights and trends applicable to the SME context. The results of this mapping revealed several digital maturity models that help understand how digitized the companies' processes are. It also highlights recurring challenges in implementing new technologies, such as lack of infrastructure, resistance to change, lack of skilled personnel, and lack of resources. Regarding benefits, the results show that the adoption of Industry 4.0 practices leads to increased product quality, production efficiency, greater customer engagement, and overall profit growth for the company.

The interviews with 10 food sector companies played a crucial role in shaping the guide, offering a practical perspective on the needs and challenges faced by these businesses. The interview data helped identify gaps between existing practices and the demands of digital transformation.

3.4 Research limitations

Although many food sector companies located within the research area were invited to participate, few responded positively. Additionally, among those that agreed to participate, not all answered every question in the questionnaire, limiting the information collected and potentially not representing the complexity and diversity in the adoption of digital transformation. The companies that did not answer all the questions were not discarded, as the sample size was small, and even though some questions were ignored, the responses obtained were important for data analysis. Furthermore, the participating companies are located in the Higienópolis neighborhood of São Paulo, near the university, and therefore may not represent the reality of Brazil, which is a highly diverse country. São Paulo is a city that hosts a significant portion of Brazilian industries. The research is also limited to a specific sector, the food industry. Moreover, the recommendations indicated were not validated. Furthermore, the responses may be subject to biases, as companies might have answered in a way they perceived as more favorable or may have underestimated or overestimated their practices and perceptions.

4 Results

The purpose of this research is to develop a practical and semi-comprehensive guide based on the responses obtained through interviews conducted with SMEs and the literature mapping carried out in the first part of this

thesis. The objective is to assist them in improving their operations in response to the challenges and opportunities presented within the context of Industry 4.0.

4.1 Interview results

Below are the most relevant data collected from interviews with SMEs in the food sector. These data highlight both the challenges faced by these companies and the opportunities for intervention and adoption of Industry 4.0 practices and processes.

Figure 1 illustrates the intention of technology adoption in businesses and the interest in implementing new technologies. SMEs show particular interest in widely adopted solutions such as barcodes (QR codes), pagers, and social media, as well as emerging technologies like artificial intelligence and robotics.

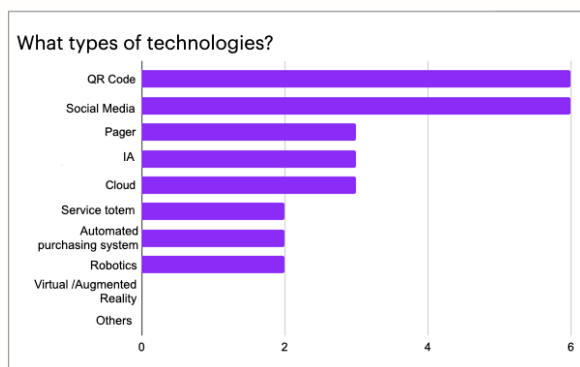


Fig. 1. Intended technologies.

The main barriers identified by the interviewed SMEs in adopting technology, as indicated in Figure 2, highlight the need for easier and more accessible solutions to reduce financial and time-related costs. The key challenges include a lack of adequate training, unfamiliarity with new technologies, and the expectation of high financial expenses, among others.

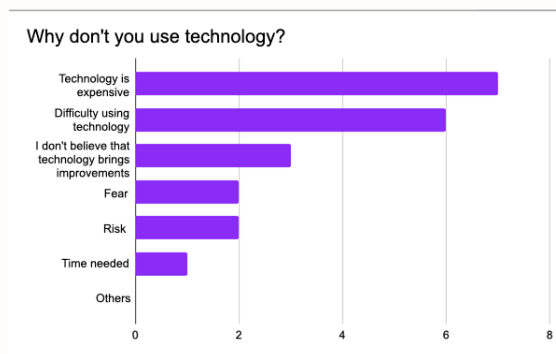


Fig. 2. Reason why SMEs do not implement new technologies.

Figures 3 and 4 demonstrate the underutilization of digitalization in business communication. The majority of interviewed SMEs lack digital marketing and promotional strategies. However, most of them consider themselves competitive in their respective markets, as indicated in Figure 5. This disparity reveals a contradiction between how these companies perceive their competitiveness and their actual potential; establishing an online presence through social media and digital marketing strategies could be a significant differentiator in further enhancing their market position.

Does the company use marketing actions to publicize news?
 10 answers

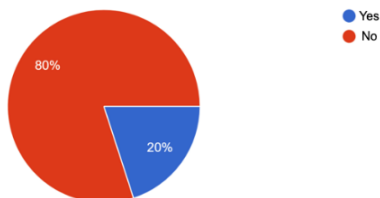


Fig. 3. Use of marketing actions.

Your company uses digital means to share information and communicate with its customers
 10 answers

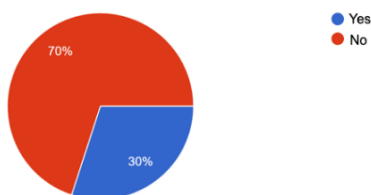


Fig. 4. Digital communication with customers.

Does the company seek to remain competitive in relation to its competitors?
 10 answers

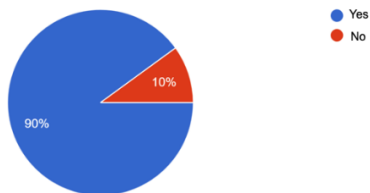


Fig. 5. Self-assessment of SME's competitiveness.

Of the participating companies, only 40% evaluate the impact of technology on their business, as shown in Figure 6. These companies report satisfaction with the results of technology use, recognizing the improvements that technology has brought to their activities.

Do you have any way of evaluating the results of using technology in your business?
 10 answers

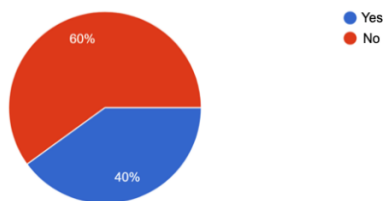


Fig. 6. Responses to assess the use of technology.

Based on the data collected, this study seeks to understand the relationship between the interviewed companies and Industry 4.0, as well as to identify recommendations that can help them adopt digital transformation strategies efficiently and effectively. The goal of this guide is to enable SMEs to develop a personalized action plan to enhance their digital maturity, allowing them to compete more efficiently and sustainably in an increasingly automated and digitalized environment.

5 Recommendation guide for food industry SMEs in adopting Industry 4.0 practices

This guide aims to assist SMEs seeking to drive digital transformation in their businesses by compiling recommendations for practices, processes, and technologies. It is divided into five main topics, each containing items related to technologies and/or enabling practices of Industry 4.0.

These topics will be presented in the following subsections. Based on these items, SMEs can be guided in adopting technologies for transformation into Industry 4.0.

5.1 Technological infrastructure, tools, and applications

Currently, there is a wide range of technological tools that enable digital transformation, support business operations, and can be easily integrated into companies. These technologies stand out primarily for facilitating data sharing, automating processes, and promoting business growth.

Table 1 presents a list of technologies that offer quick integration and are commonly used in the daily operations of food industry SMEs. A tutorial is provided, explaining their functionality, potential challenges, and implementation costs values referring to the first semester of 2024.

Table 1. Integration technologies.

Technology	Costs	Difficult	Tutorial
QrCode	1. Implementation cost: -QR code generator: Free - BRL 50 -Printing: BRL 0.10 - BRL 1 per unit of paper 2. Cost of training and adaptation: quick and simple, as it only involves creating and using the QR codes.	Low	A QR Code generator can easily be found on the internet, such as QR Code Generator. It can be used to direct you to a menu or a tab with more information about the company and thus increase engagement. Then just print it out and use it
Pager	1. Implementation cost: -Pager system: BRL500- BRL2000 2. Cost of training and adaptation: it can take a few weeks for employees to get used to it	Medium	Buy a pager system and follow the instructions in the manual to set up the pagers and base correctly
Service totem	1. Implementation cost: -Totem: BRL 2000- BRL 8000 2. Cost of training and adaptation: it can take up to a month for complete integration into the workflow	High	Choose a model that meets your needs, as you will need more in-depth knowledge of technology to configure the software. Some totems come with a manual. To use a service totem, you need an internet connection
Digital menu	1. Cost of implementation: -creation of the menu: free- BRL500	Medium	You can use menu services (Menudino, iMenu360) or create your own pdf. Then



5.2 Management and organizational culture

- Use of messaging apps for communication between employees and sharing relevant information. Apps like WhatsApp and Telegram can be used for quick and informal communication, facilitating information exchange and task organization, while email can be used for sending important documents and recording official conversations.
- Implementation of a management system, preferably an ERP system, to control sales and business activities. This system integrates all company areas, enabling greater centralization of information (sales and inventory tracking, order management, and detailed reporting), improving decision-making.
- Hiring IT consultancy services to implement, maintain, and evaluate the use of technologies, especially management systems. IT consultants can provide continuous support to ensure technologies function efficiently, measure performance, and suggest improvements.

- Inventory control systems (BLing, Tiny, IBM Cloud Pak)
- Order distribution (Self-service kiosks, queue management systems, order apps)
- Delivery services using apps (iFood, Rappi, etc.)
- Product storage solutions

5.4 Strategy and innovation

This section discusses how businesses can use the information obtained through digitalization to analyze key areas and improve performance. It is recommended that, alongside digital transformation efforts, businesses start with low-cost technologies and practices such as social media apps, QR codes, delivery apps, and inventory systems. Additionally, companies should evaluate the impact of these changes by asking:

- How well have employees adapted to new technologies?
- Have there been gains or losses in information processing due to the introduction of new technology/practices?
- Has the new technology/practice provided new insights into customer behavior?
- Has the new technology/practice provided new insights into market trends?
- Based on the data collected, can process improvements be suggested?

By addressing these questions and conducting a critical analysis of the transformation process, companies can develop strategies to maintain and improve their operations.

5.5 Market and customers

To remain competitive, small and medium-sized enterprises (SMEs) in the food industry must adopt a proactive digital approach. Digital platforms help businesses connect with customers, understand consumption trends, and expand their reach. Social media platforms like Instagram, TikTok, and Facebook allow businesses to engage with consumers, share relevant content, and creatively promote products. Maintaining a strong online presence not only attracts new customers but also strengthens loyalty among existing ones.

Some useful technologies and tools for SMEs include:

- Google Analytics (Free tool for analyzing website traffic)
- Google Alerts (Allows businesses to set up alerts to monitor online mentions of their brand, products, or competitors)
- Simple Data Analysis Tools (Spreadsheets like Microsoft Excel or Google Sheets, or tools like OpenRefine)
- Manual Competitor Monitoring (A viable option for small businesses, involving active monitoring of competitor activities on social media, websites, and other online channels)

6 Discussion

Insights gathered from SMEs in the food industry reveal a general awareness of the need for digital transformation, though barriers such as cost, lack of knowledge, fear, and time constraints lead to hesitation. This resistance, often referred to as "decision-maker myopia," is common among SMEs, as noted by [10] in his analysis of technological adoption challenges. The low adoption of automated customer service tools and digital platforms highlights a gap in the implementation of solutions that could improve operational efficiency and customer experience, confirming the findings of [8].

Despite the fact that half of the surveyed companies have already automated some processes, challenges remain in process automation and digital integration. This underscores the need for investment in IT consulting services and the development of metrics to evaluate technology return on investment.

The growing interest in innovation and technology adoption, including QR codes and process automation, indicates a favorable environment for the implementation of more advanced digital solutions. [10] suggests that market competition also plays a significant role in pushing SMEs toward innovation, as businesses feel compelled to adopt new technologies when their competitors and suppliers do.

The recommendation guide presented in this study includes technologies such as QR codes and self-service kiosks, which improve customer interaction and accessibility. The use of social media for internal and external

communication has also been effective in measuring customer satisfaction, as demonstrated in the hospitality industry by [7]. Therefore, leveraging these tools is essential for refining marketing strategies.

The interview results highlight the urgent need to overcome perceived barriers and fully embrace digital transformation in the food sector. The analysis of collected data suggests that a strong commitment to innovation, strategic technology investments, and a company culture that values continuous learning are essential. As [7] noted, businesses that successfully navigate this transition will be better positioned to meet the evolving demands of the market, benefiting from increased revenue, reduced operational costs, more efficient daily processes, higher employee productivity, and enhanced customer support.

7 Conclusion and future work

This study aimed to identify and compile a set of best practices and recommendations to assist SMEs in their digital transformation journey and in adopting Industry 4.0 practices. A systematic mapping technique was employed to establish a theoretical foundation, revealing that there is still a lack of relevant publications on this subject in both academic (white literature) and non-commercial sources (gray literature).

To understand the relationship between SMEs and Industry 4.0, a set of common digital maturity practices was established. These recommendations were then used to design a questionnaire, which was applied in interviews with SMEs in the food sector, leading to the creation of a recommendation guide. By focusing on a specific sector and geographic location, the study was able to generate tailored suggestions that are more applicable to food industry businesses. However, since digital maturity challenges are widespread among SMEs, the recommendations may not be universally valid for other industries.

As future work, it is suggested that the proposed recommendation guide be validated and subsequently implemented in SMEs within the food industry in central São Paulo. Additionally, the guide should be adapted and validated for use in SMEs from other sectors, thereby expanding its applicability.

Acknowledgement. This work was supported by MackPesquisa (Mackenzie Research and Innovation Fund [Fundo Mackenzie de Pesquisa e Inovação]), Project number 221006.

References

1. Palmeira, J., Coelho, G., Carvalho, A., Carvalhal, P. and Cardoso, P.: "Migrating legacy production lines into an Industry 4.0 ecosystem," *2022 IEEE 20th International Conference on Industrial Informatics (INDIN)*, Perth, Australia, 2022, pp. 429-434, doi: 10.1109/INDIN51773.2022.9976084.
2. Passos, L. H. S.: Indústria 4.0: Fundamentos e Principais Impactos na Economia Brasileira. *Revista de Administração e Negócios da Amazônia*, v. 12, pp. 53 (2020). <https://doi.org/10.18361/2176-8366/rara.v12n2p53-63>
3. Brasil: Agenda Brasileira para a Indústria 4.0. Brasília (2018). <https://www.industria40.ind.br/>, last accessed 2023/04/20
4. Sebrae: Micro e pequenas empresas geram 27% do PIB do Brasil (2021), <https://sebrae.com.br/sites/PortalSebrae/ufs/mt/noticias/micro-e-pequenas-empresas-geram-27-do-pib-do-brasil.ad0fc70646467410VgnVCM2000003c74010aRCRD>, last accessed 2023/05/20
5. Cristiano, G.B., Tanikawa, F., García, H.L., Eliseo, M.A.: The Adoption of Industry 4.0 Practices for Small and Medium-Sized Companies: A Systematic Mapping Study. In: Ruiz, P.H., Agredo-Delgado, V., Mon, A. (eds) *Human-Computer Interaction. HCI-COLLAB 2023. Communications in Computer and Information Science*, vol 1877. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-57982-0_23
6. Santos, A. O.: Indústria 4.0 e suas interfaces em uma empresa de pequeno porte no interior da Bahia. Estudo de Caso. USSC, Santa Catarina (2021). <https://repositorio.animaeducacao.com.br/handle/ANIMA/17753>.
7. Gaspar, L. S.: Transformação digital no setor hoteleiro de São Luís – MA: entraves e desafios para gestores de pequenas empresas (2022), <https://monografias.ufma.br>, last accessed 2025/03/15.
8. Machado, D. A. De S.: Transformação digital de modelo de negócio tradicional de lavanderias: uma proposta de framework. In: Universidade Nove de Julho. Repositorio Nove de Julho (2021). <https://bibliotecatede.uninove.br/handle/tede/2897>
9. Eliseo, M. A. et al. An Overview of Brazilian Companies on the Adoption of Industry 4.0 Practices. In: Agredo-Delgado, V., Ruiz, P.H., Correa-Madrigal, O. (eds) *Human-Computer Interaction. HCI-COLLAB 2022*.

- Communications in Computer and Information Science, vol 1707, pp 15-27 Springer, Cham (2023).
https://doi.org/10.1007/978-3-031-24709-5_2
10. Fassini, R.: Inovação em modelo de negócios através do uso de tecnologias da indústria 4.0 em pequenas e médias empresas (2021), <https://tede2.pucrs.br/tede2/handle/tede/9893>, last accessed 2025/03/15
 11. Brasil, Ministério da Ciência, Tecnologia, Inovações e Comunicações: Guia de Referência para atuação em Gestão da Inovação com foco em Transformação Digital, 2023, <https://www.portaldaindustria.com.br/publicacoes/2023/1/guia-de-referencia-para-atuacao-em-gestao-da-inovacao-com-foco-em-transformacao-digital/>, last accessed 2023/05/28
 12. ABDI, Agência Brasileira de Desenvolvimento Industrial, Aplicação de tecnologias e inserção digital (2021). https://api.abdi.com.br/file-manager/upload/files/Instituto_FSB_Pesquisa_-_ABDI_-_Conectividade_VF_2_1.pdf.



4º WORKSHOP SOBRE TRABAJO COLABORATIVO Y APRENDIZAJE COLABORATIVO

Prototipo de Aplicación Móvil para la Detección de Tendencias Auto - Lesivas en Estudiantes de Medicina

José G. González Rodarte ¹ [0009-0005-5544-8755], Esmeralda Carrillo Chávez ¹ [0009-0005-3908-949X],
Jaime Mercado Reyna ¹ [0009-0009-3617-1243], Huizilopoztli Luna-García ¹ [0000-0001-5714-7482]
¹ Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX), Unidad Académica de
Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Zacatecas 98000, México

Resumen. Actualmente se ha observado un incremento en padecimientos mentales, promoviendo el riesgo de presentar tendencia auto-lesivas. La población de estudiantes de la licenciatura de medicina de la Universidad Autónoma de Zacatecas no es la excepción, debido a la demanda mental y física que conlleva. En respuesta a esta problemática se realizó el siguiente estudio con el objetivo de desarrollar un prototipo de aplicación móvil nombrado "SelfMind" cuya finalidad es la recolección de datos mentales de estudiantes de medicina y emplearlos para detectar tendencias autolesivas de manera anticipada. Para el desarrollo del prototipo "SelfMind" se implementó la metodología "Diseño Centrado en el Usuario" siguiendo el estándar ISO-9241-210. Los datos cualitativos se recabaron utilizando la técnica "Persona" con una muestra de 10 estudiantes y 2 psicólogos profesionales, posteriormente la evaluación de las versiones del prototipo se realizaron con 5 estudiantes y 2 psicólogos. Los resultados obtenidos se utilizaron para modificar el prototipo. Como resultado se obtuvo un prototipo de software con funcionalidades implementadas permitiendo un mejor acercamiento a una versión final. El prototipo consiguió 100% de tareas completadas exitosamente, promedio de tiempo por tarea de 22.8 segundos, promedio de 4.9 clics por tarea y puntaje promedio en el instrumento IBM Questionnaire for User interaction Satisfaction de 3.947. Se concluyó que el uso de herramientas y metodologías de Diseño Centrado en el Usuario contribuyó a la usabilidad en sus 3 características: eficiencia, eficacia y satisfacción. Esta contribución se puede observar en la evolución del prototipo y en las métricas utilizadas.

Keywords: Palabras clave: Diseño Centrado en el Usuario, Aplicación Móvil, Tendencias Auto-Lesivas.

1 Introducción

La presencia de tendencias auto-lesivas en una persona es el resultado de un conjunto de síntomas que se dividen en 5 etapas; ideación auto-lesiva pasiva, contemplación activa, planeación, ejecución de intentos auto-lesivos y la práctica auto-lesiva consumada [1]. La Organización de las Naciones Unidas reporta más de 97.000 casos de autolesiones con resultados mortales en el continente americano en 2019, de los cuales el 79% son hombres [2]. En México, las defunciones que resultan de estas conductas han aumentado considerablemente en el país. Según datos de la Encuesta Nacional de Bienestar Autorreportado, en 2017 la tasa de comportamientos autodestructivos fue de 5.3 casos por cada 100 mil habitantes y para el 2022 paso a 6.3 casos respecto a 2017, sin embargo, en 2022 disminuir el número de bajas, registrando un total de 213 casos [3]. Esto es equivalente a 1629 fallecimientos más en 2022 respecto a 2017, además el 13 % de mujeres de 18 años o más notificaron sentirse en depresión la mitad o incluso más días de la semana anterior a la encuesta en contraste con 9.1% en el caso de los hombres que reportó la misma condición [4].

Estos problemas parecen tener una mayor incidencia dentro de la población de Médicos e incluso presentándose desde su formación de pregrado en medicina [5]. Esta clase de conductas están muy relacionadas con índices de estrés, así como de depresión [6]; condiciones presentes en la formación de Medicina Humana y que afectan a una gran parte de la población de estudiantes [7], por lo cual es necesario promover espacios de discusión sobre estos temas y se tomen acciones para la intervención temprana en estudiantes en alto riesgo de conductas auto-lesivas [6].

Gracias al avance tecnológico de en los últimos tiempos, se ha tratado de abordar este tipo de problemáticas referentes a la salud mental, dando lugar a nuevas tecnologías conocidas como e-mental health, cuyo propósito

es ayudar al diagnóstico y servir como herramienta de apoyo para el tratamiento de estas patologías [8]. No obstante, se presentan varios retos al momento de la implementación de estas herramientas.

El problema principal al que se enfrentan es la falta de adopción por parte de los usuarios, lo cual se define como un rotundo fracaso [8]. Para mitigar este problema, se ha implementado la metodología de Diseño Centrado en el Usuario de modo que los usuarios se involucren más en el diseño de estas aplicaciones fomentando su usabilidad. Esta metodología ayuda a los desarrolladores a entender mejor las necesidades de los usuarios generando mejores resultados y mayor grado de aceptación [9].

El potencial que tienen estas herramientas de e-mental health es bastante amplio [8], por lo cual la Universidad Autónoma de Zacatecas ha realizado una primera implementación de una herramienta de esta naturaleza llamada "Mentali" con el propósito de facilitar a los psicólogos universitarios identificar problemas de ansiedad y depresión en los estudiantes de la facultad [10]. A pesar de la existencia de esta herramienta se busca desarrollar una diferente, de modo que implemente un modelo de Aprendizaje Automático para la detección de tendencias auto-lesivas en los estudiantes, además de implementar la metodología de Diseño Centrado en el Usuario para promover la usabilidad.

En el presente artículo se detalla el proceso realizado para el diseño y desarrollo de la interfaz de esta nueva herramienta.

2 Metodologías

2.1 Diseño centrado en el usuario

Para el desarrollo de este prototipo se utilizó la metodología propuesta por Donald Norman, Diseño Centrado en el Usuario, la cual se define como una metodología interactiva de desarrollo donde el usuario influye en el producto final, priorizando las necesidades e intereses del usuario, haciendo énfasis en que los productos sean utilizables y comprensibles [11]. Esta metodología es primordial para garantizar una aceptación del producto por parte de los usuarios de modo que se sientan cómodos y satisfechos con el resultado final.

Para su aplicación se siguió el estándar ISO 9241-210 en su versión de 2019. Este estándar establece cuatro principios; 1) Involucrar activamente a los usuarios y entender los requisitos y las tareas a realizar, 2) correspondencia adecuada entre las funcionalidades de la tecnología y los usuarios, 3) iteración de soluciones de diseño y 4) diseño multidisciplinar [12]. De igual manera el estándar divide el proceso en cuatro fases; entender y especificar el contexto de uso, especificar los requerimientos de usuario, producir soluciones de diseño en base a los requerimientos y evaluación del diseño contra los requerimientos (ver Fig. 1).

Fase 1: Entender y especificar el contexto de uso. En esta etapa se caracteriza a los usuarios y el contexto en el que el producto será usado, información que se tiene que tomar en cuenta para el desarrollo de un buen diseño.

Para este estudio en particular se realizó la técnica "Persona", la cual permite identificar las necesidades, deseos, experiencias, metas y obstáculos de los usuarios. Para la obtención de los diferentes perfiles Persona se solicitó la ayuda de 10 estudiantes de medicina humana de la Universidad Autónoma de Zacatecas y 2 profesionales del área de Psicología. Posteriormente se realizó una serie de entrevistas particulares para obtener descripciones de mapas mentales referentes a herramientas e-mental health y detección de objetos.



Fig. 1. Fases del diseño centrado en el usuario (elaboración propia inspirada en el estándar ISO 9241-210 [12]).

Fase 2: Especificación de los requerimientos. Con la información recolectada de la fase anterior se procedió a la especificación de los requerimientos. Para esta fase se siguió el estándar IEEE 830 el cual nos proporciona una guía de especificación de requerimientos de software. Este estándar establece un conjunto de prácticas recomendadas para documentar los requerimientos de un sistema de software, asegurando que sean claros, completos y verificables [13]. Además, para complementar esta fase se utilizó la técnica de Historias de Usuario, utilizada en el desarrollo ágil de software para capturar requerimientos funcionales desde la perspectiva del usuario final [14].

Fase 3: Producir soluciones de diseño. Con los requerimientos ya definidos se procedió al desarrollo de una solución del diseño tomando en consideración la información recopilada de las fases anteriores. Esta etapa constó de 3 iteraciones y con la retroalimentación obtenida en la fase 4 (evaluación del diseño contra los requerimientos) el prototipo se modificó para satisfacer los requerimientos e interacciones con los usuarios. Para el desarrollo de la solución de diseño se optó por producir un prototipo de software con el fin de añadirle funcionalidad, obteniendo un prototipo cercano a lo que sería el producto final. El prototipo se desarrolló para un sistema operativo Android 10 y posteriores utilizando el lenguaje multiplataforma Java.

Fase 4: Evaluación del diseño contra los requerimientos. A diferencia de otras metodologías, en Diseño Centrado en el Usuario se involucra de una manera más directa los usuarios finales del producto para evaluarlo, esto con el fin de abordarlo desde una perspectiva de usabilidad. Esta evaluación permite la mejora de la solución de diseño descubriendo puntos fuertes así como puntos débiles.

Para el presente estudio, se recurrió al estándar ISO 9241-11 en su versión de 2018, en el cual la usabilidad se centra en tres aspectos; eficiencia, eficacia y satisfacción [16]. Para la evaluación de estos 3 aspectos, se llevaron a cabo pruebas de usabilidad moderadas, en las cuales un facilitador guía a los participantes mientras interactúan con un producto, permitiendo obtener información detallada sobre la experiencia de usuario, sus dificultades y comportamientos al usar el producto [15]. Específicamente las pruebas fueron el registro de clics para evaluar la eficiencia y tasa de éxito para la eficacia. Durante las mismas se indicó a los usuarios una serie de tareas a realizar por medio del prototipo, se registró si el usuario finaliza con éxito la tarea, el tiempo en el que la concretó, el número de clics que necesitó y si presentó problemas de algún tipo. A cada tarea se le asignó un número de clics y tiempo máximos para concretarse. Una prueba se consideraba exitosa si el usuario finalizaba con éxito la tarea sin presentar algún tipo de problema y no excedía el número de clics y tiempo máximos. En el aspecto de satisfacción se implementó el Questionnaire for User Interaction Satisfaction de IBM que consta de 19 preguntas evaluadas en una escala de Likert, que van desde "Muy insatisfecho" hasta "Muy satisfecho"[17].

Estas evaluaciones se llevaron a cabo con 5 estudiantes voluntarios de medicina humana de la Universidad Autónoma de Zacatecas, así como 2 profesionales de Psicología.

3 Resultados obtenidos

3.1 Entender y especificar el contexto de uso

En esta primera fase se recopilamos aspectos clave de los participantes estudiantes: edad, semestre que cursan, personalidad, metas, obstáculos, motivaciones, deseos y experiencias con enfermedades mentales, así como herramientas de software de naturaleza e-mental health haciendo énfasis en la herramienta “Mentali”. Una vez recopilada la información de los 10 estudiantes se prosiguió a realizar los perfiles para cada uno de los participantes usando (ver Fig. 2). En el caso de los 2 profesionales en Psicología se realizaron sesiones individuales en las cuales se llevaron a cabo entrevistas no estructuradas donde se les cuestionó sobre sus metas, motivaciones, personalidad, obstáculos, experiencias con herramientas e-mental health y sus opiniones. Posteriormente se crearon los perfiles Persona (ver Fig. 2), el cual es diferente a los perfiles “Persona” de los estudiantes

3.2 Especificación de requerimientos

Con la información recopilada se identificaron 2 perfiles de usuario: estudiante que realizará registros de su estado de ánimo y salud mental en la aplicación y psicólogo universitario encargado de monitorear a los usuarios de tipo Estudiante.

Los estudiantes tenían muy poca experiencia con herramientas e-mental health, además de entre los 10 solo uno tenía conocimiento sobre la herramienta “Mentali” y confirmó haberla utilizado, pero por un periodo de tiempo bastante corto, lo que podría indicarnos que la herramienta no es muy usada en la facultad. También se identificó que los estudiantes perciben su entorno como hostil, estresante, con una presión constante y horarios atareados que representan una carga excesiva de trabajo de modo que tienen muy poco tiempo para realizar sus actividades. Por parte de los profesionales en Psicología, al igual que los estudiantes, reportaron tener muy poca experiencia con herramientas e-mental health, pero no descartan el uso de estas, especialmente si les ayuda con la administración de sus pacientes. También recalcaron que una herramienta e-mental health no puede reemplazar la atención psicológica de un profesional, por lo que esta debe funcionar como un apoyo. Por último, ambos grupos presentaban preocupaciones en cuanto a temas de seguridad de las herramientas e-mental.

(Nombre)			
Metas (Metas u objetivos de la persona)	Motivación (¿Qué impulsa a la persona a conseguir sus metas y deseos?)	Metas (Metas u objetivos de la persona)	Motivación (¿Qué impulsa a la persona a conseguir sus metas y deseos?)
(Fotografía)		(Fotografía)	
Obstáculos (¿Obstáculos que se enfrenta la persona para tratar o solucionar su salud mental?)	Deseos (¿Cuáles son los intereses y deseos de la persona referente a su salud mental y su ambiente escolar?)	Obstáculos (¿Obstáculos que se enfrenta la persona para tratar o ayudar a sus pacientes?)	Deseos (¿Cuáles son los intereses y deseos de la persona referente al estado de salud mental y su profesión?)
(Frase representativa)		(Frase representativa)	
Edad (Años cumplidos de la persona)	Observaciones (Cualquier información relevante sobre la persona, como sensaciones y percepción de su ambiente, relación entre sus compañeros y docentes)	Edad (Años cumplidos de la persona)	BIO (Biografía corta de la persona así como una pequeña rutina diaria)
Semestre (Semestre que cursa actualmente la persona)	Rendimiento (Percepción de las personas de su rendimiento académico)		Experiencia (Experiencia de la persona con software de carácter e-mental health)
	☐ Mala ☐ Regular ☒ Bueno ☐ Excelente		
Personalidad (Características de la personalidad de la persona)	Historia Salud Mental (Prácticas relacionadas a la salud mental de la persona)	Personalidad (Características de la personalidad de la persona)	

Fig. 2. Plantilla persona tipo estudiante y tipo psicólogo (elaboración propia).

Con la información obtenida se pudieron concretar necesidades de los 2 tipos de usuario y definir funcionalidades para cada perfil de usuario y atributos del prototipo. Como funcionalidades para el perfil de usuario Estudiante se identificaron: la creación y vista de bitácoras donde los usuarios pudieran expresarse, realización de cuestionario centrado en la detección de trastornos como ansiedad y depresión para la recolección de datos y retroalimentación al usuario. Por parte del perfil de usuario Psicólogo se identificaron: la lectura de las condiciones de un usuario Estudiante en específico y visualización de los usuarios Estudiante.

3.3 Producción de soluciones de diseño y evaluación

Estas dos fases son iterativas por lo que se realizaron 3 iteraciones de la fase de producción de soluciones de diseño y 2 de evaluación. Como se mencionó anteriormente la evaluación se centró en la usabilidad haciendo énfasis en los aspectos de eficiencia, eficacia y satisfacción según el estándar ISO 9241-11. La eficiencia se evaluó con la técnica de Registro de clics, la eficacia con Tasa de éxito y finalmente la satisfacción con el instrumento Questionnaire for User Interaction Satisfaction de IBM.

Producción de soluciones de diseño: primera iteración. Para esta primera iteración se desarrolló un prototipo bastante básico, el cual incluía las funcionalidades principales encontradas durante la fase de requerimientos. Como color principal de este primer desarrollo de solución de diseño se utilizó un color en tono púrpura (específicamente Purple 500) esto debido a que, según la guía para el diseño visual Material Design de Google es un color vibrante y funciona muy bien en interfaces digitales además de ser lo suficientemente neutral para ser una base adaptable [18].

Este prototipo se dividió en 2 secciones, una para los usuarios de tipo Estudiante y la siguiente para los usuarios de tipo Psicólogo. La sección de Estudiante se compone de 4 interfaces:

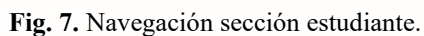
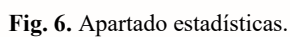
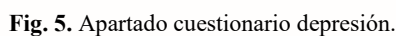
- Menú. Un menú simple donde el usuario selecciona la actividad o tarea que desea realizar (ver Fig. 3).
- Creación de Bitácoras. En esta interfaz el usuario puede crear una bitácora con un estado de ánimo y escribiendo lo que desee (ver Fig. 4). Para las opciones de estado de ánimo, se usaron aquellas definidas en la aplicación móvil "Mentali" [10].
- Cuestionario Depresión. El usuario podrá responder una serie de preguntas orientadas a recabar información respecto a su estado mental (ver Fig. 5). El cuestionario implementado es el utilizado dentro de la facultad de medicina humana de la universidad y en la aplicación "Mentali" [10]. Para interactuar con todo el contenido dentro de esta interfaz simplemente se desliza la pantalla de arriba hacia abajo.
- Estadísticas. En este apartado el usuario podrá ver un resumen de su estado mental en base a las bitácoras creadas y los cuestionarios realizados (ver Fig. 6). En esta interfaz la interacción es exactamente igual a la interfaz de Cuestionario Depresión.



Fig. 3. Menú estudiante.



Fig. 4. Apartado creación de bitácoras.



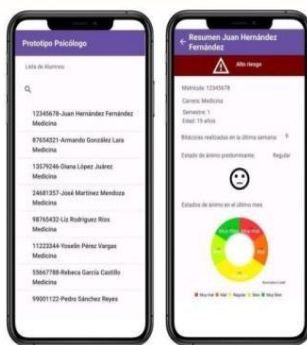


Fig. 8. Sección psicólogo.



Fig. 9. Interfaz resumen estudiante.

Para esta sección del prototipo, se implementó la arquitectura de navegación basada en botones (Button Navigation) donde los botones ubicados en una pantalla principal permiten a los usuarios acceder a diferentes funcionalidades donde cada botón está claramente etiquetado y vinculado a una vista específica (ver Fig. 7).

La sección orientada al usuario Psicólogo (ver Fig. 8) se conforma por una lista de usuarios Alumnos con un buscador integrado de modo que se pueda buscar un alumno en específico y cuando se seleccione cualquiera de los objetos en la lista aparezca el resumen más detallado. La navegación en la interfaz destinada al resumen de un alumno igualmente se desliza la pantalla de arriba hacia abajo (ver Fig. 9).

Evaluación del Prototipo: primera iteración. Todas las tareas proporcionadas a los usuarios estudiantes para concretarlas por medio del prototipo las finalizaron correctamente y a tiempo, sin embargo 3 de los 5 participantes excedieron el número de clics máximo y presentaron problemas de navegación, dificultades para encontrar datos específicos, complicaciones para lectura del contenido del prototipo (tamaño y fuente de la letra) así como falta de instrucciones en la aplicación. En el cuestionario de satisfacción los estudiantes reportaron problemas con la navegación, consecuentemente estos contratiempos producían problemas adicionales a los usuarios al usar el prototipo como la localización de información. Sumado a eso los usuarios reportaron inconformidad en aspectos de seguridad e instrucciones.

En cuanto a la evaluación de los 2 usuarios Psicólogos, todas las tareas que se asignaron las finalizaron, pero excedían el número de clics máximo, así como el tiempo por un margen considerable. Se destacaron problemas de navegación, de lectura y búsqueda de información y señalaron que el color principal no era del todo agradable, ya que existen otros colores que promueven de una mejor manera la relajación como tonalidades de verde o azul. Respecto a satisfacción remarcaron malestar referente a la navegación, localización y lectura de información y seguridad. Finalmente compartieron mejoras que se podrían implementar, especialmente en la interfaz de lista de alumnos, ya que exhiben grandes dificultades para localizar a un usuario registro específico.

Para esta primera evaluación se obtuvieron un 100 % de las tareas completadas, un número de clics por tarea de 6.86, un tiempo de 55.44 segundos por tarea y un promedio de 2.47 en el cuestionario de satisfacción.

Producción de soluciones de diseño: segunda iteración. Con los resultados y comentarios de los usuarios en la primera evaluación se modificó el prototipo inicial. De manera general se aplicaron las siguientes modificaciones:

- Implementación de un login. El propósito es mejorar los problemas de seguridad destacados por los usuarios (ver Fig. 10).
- Arquitectura de navegación. Se implementó una arquitectura Bottom Navigation, similar a aplicaciones de redes sociales con el propósito de que los usuarios se sientan más familiarizados con el prototipo (ver Fig. 11).
- Texto y color. Como color principal se optó por una tonalidad verde sugerida por los 2 profesionales. Así mismo se modificó la fuente y de tamaño de letra.
- Interacción en las interfaces de Estadísticas y Resumen Alumno. Se reemplazó el enfoque de deslizamiento de arriba abajo por un enfoque lateral con el fin de facilitar la búsqueda y lectura de información para los usuarios(ver Fig. 13 y Fig. 16)
- Correspondiente a la sección de usuarios Estudiantes (ver Fig. 12) se realizaron las siguientes mejoras:
- Campo de Título bitácora. Para la creación de una nueva bitácora, se agregó el campo de Título con el propósito de que el usuario pueda buscar una bitácora en específico y facilitar la búsqueda de información (ver Fig. 14).
- Apartado de Lista bitácoras. Se añadió el apartado de Lista bitácoras donde se podrá buscar una bitácora específica, ver sus características y eliminarla(ver Fig. 15).



Fig. 10. Login

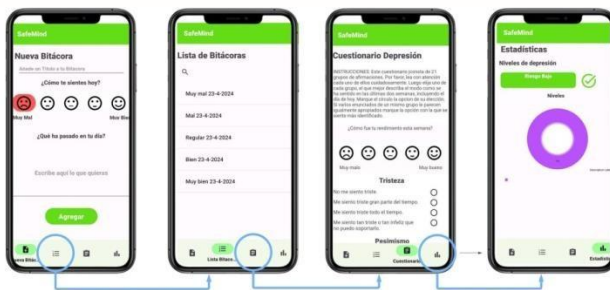


Fig. 11. Navegación prototipo 2.

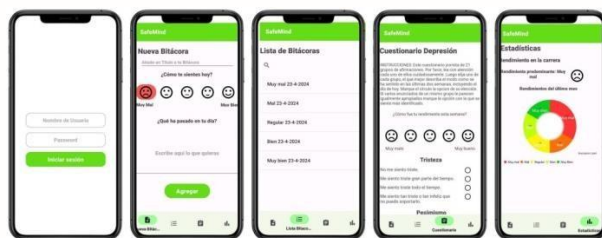


Fig. 12. Sección estudiante prototipo 2.



Fig. 13. Navegación de Estadísticas prototipo 2.



Fig. 14. Nueva bitácora prototipo 2.



Fig. 15. Lista bitácoras

Fig. 16. Navegación resumen estudiante prototipo 2



Fig. 17. Sección psicólogo prototipo 2.

Por otro lado, en la sección de usuarios Psicólogo (ver Fig. 17) se dotó al prototipo con las siguientes características:

- Mejora en Lista de Alumnos. Se realizaron modificaciones al formato para ayudar a la visibilidad de los datos. Adicionalmente se añadió el estado de riesgo a cada ítem de la lista con la finalidad de que el usuario siempre disponga de este dato en primera instancia sin la necesidad de acceder a otra interfaz (ver Fig. 18)
- Interfaz Alertas. Implementación de una interfaz en la cual aparecen aquellos usuarios Alumnos con un grado de Riesgo Alto cuyo aporte es que él se identifique de manera más eficaz y rápida aquellos Alumnos que requieran una atención inmediata (ver Fig. 19).



Fig. 18. Lista Estudiantes prototipo 2.



Fig. 19. Alertas prototipo 2.

Evaluación del prototipo: segunda iteración. Durante esta evaluación los estudiantes completaron la mayoría de las tareas sin exceder el número de clics máximos y sin sobrepasar el tiempo límite. Entre las mejoras, los participantes reportaron una navegación más cómoda en las interfaces del prototipo y una mayor facilidad en la visualización del texto.

Sin embargo, en la interfaz de Estadísticas presentaron grandes dificultades para adoptar el nuevo enfoque de navegación lateral del contenido, consecuentemente desencadenó un excedente exagerado de número de clics y tiempo máximos. Incluso 4 de ellos no finalizaron con éxito la tarea asignada. Remarcaron una falta de instrucciones o indicios sobre el funcionamiento de la nueva interacción con la interfaz. Paralelamente, los participantes presentaron problemas al identificar los botones de regreso e informaron obstáculos en las interfaces de Nueva Bitácora y Cuestionario Depresión, esto a causa de se dejaron campos sin responder por lo que se incrementó el número de clics. En lo que corresponde a satisfacción, los estudiantes expresaron una conformidad con la seguridad. Los contratiempos ocurridos en la interfaz de Estadísticas se reflejaron como inconformidad en el aspecto de búsqueda y visualización de información. Por otro lado, los participantes presentaron cierto disgusto con el color principal, declararon que la tonalidad de verde elegida es muy brillante, lo causaba molestias leves al momento de utilizar el prototipo.

Referente a la evaluación realizada por los participantes Psicólogos reportaron tener grandes dificultades con la interfaz de Resumen Alumno, debido a que la forma de acceder a las diferentes gráficas de la interfaz era poco intuitiva, por lo tanto, ninguno de los dos participantes logró terminar la tarea con éxito. Sin embargo, reportaron una gran mejora en la visualización de la Lista de Alumnos. Expresaron experimentar una mejor interacción y facilidad para encontrar un usuario alumno en específico. Adicionalmente manifestaron que el tener una lista de alumnos exclusiva de riesgo alto les permite identificar con mayor practicidad aquellos alumnos que requieren atención con más urgencia. En el cuestionario de satisfacción, reportaron una mejora significativa en la interacción del prototipo, navegación y seguridad, pero presentaron disgusto con la facilidad de encontrar información, específicamente en la interfaz de resumen alumno.

Para esta segunda evaluación se obtuvieron un 91 % de las tareas completadas, un número de clics por tarea de 7.74, un tiempo de 46.23 segundos por tarea y un promedio de 3.4 en el cuestionario de satisfacción. Comparado con el prototipo anterior, se empeoró en tareas completadas y número de clics, en las métricas restantes mejoró.

Producción de soluciones de diseño: tercera iteración. Durante la segunda evaluación, los participantes presentaron dificultades en las interfaces de Estadísticas y Resumen Alumno, por esta razón se agregaron pestañas indicado donde se encuentra la información del usuario (ver Fig. 20). Para esta versión se modificó el color principal, escogiendo una tonalidad de azul grisáceo (propuesto por los participantes Psicólogos) con el fin de no provocar molestias por colores brillantes en los usuarios.

En la sección de usuarios Estudiantes (ver Fig. 21) se realizaron cambios en:

- Visualización en interfaz Lista Bitácoras. Cada ítem cuenta con el título de la bitácora, la fecha en la que se creó y el estado de ánimo asociado con el propósito de ayudar al usuario a encontrar una bitácora específica (ver Fig. 22).

- Botón de regreso. En la interfaz Detalle Bitácora, se reemplazó el botón “Listo” con una flecha hacia la izquierda para retroceder a la interfaz anterior, posteriormente se reubico el botón eliminar (ver Fig. 22).
- Visualización en la interfaz Cuestionario Depresión. El cuestionario se modificó para que cada sección fuera una lista desplegable para no saturar al usuario de información, además cuando el usuario contesta una sección esta cambiará de color indicando que fue contestada (ver Fig. 23).



Fig. 20. Navegación interfaz Estadísticas prototipo 3.

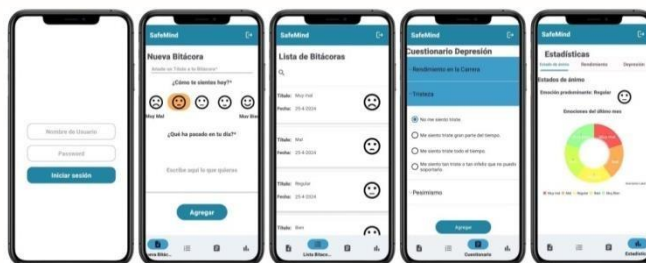


Fig. 21. Sección Estudiante prototipo 3.

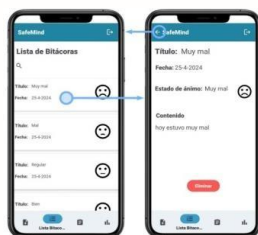


Fig. 22. Lista Bitácoras prototipo 3.



Fig. 23. Cuestionario Depresión prototipo 3

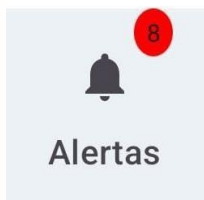


Fig. 24. Contador alertas prototipo 3



Fig. 25. Lista Estudiantes prototipo 3

Para la sección de usuarios Psicólogos (ver Fig. 24), se realizaron cambios en:

- Contador de alumnos con alto riesgo. En el ítem Alertas del menú se colocó un contador que indica al usuario el número de alertas que tiene en ese instante (ver Fig. 25)
- Botón de regreso. Se reemplazó el botón "Listo" por una flecha a la izquierda en la interfaz Resumen Alumno para regresar a la interfaz anterior (ver Fig. 25)

Evaluación del prototipo: tercera iteración. Los usuarios alumnos completaron con éxito las tareas asignadas en su totalidad sin exceder el número de clics y tiempo máximos, no presentaron problemas con el color principal ni en la búsqueda de información específica. Sin embargo, 2 de ellos notificaron que el tamaño de la fuente en algunos componentes era pequeño, por ejemplo, en las pestañas correspondientes a la interfaz Estadísticas, pero no fue impedimento para completar las tareas asignadas. En el cuestionario de satisfacción, respecto a la navegación, seguridad, interacción, búsqueda de información, los usuarios externaron satisfacción, paralelamente reflejaron problemas leves referente a la velocidad del prototipo.

Los participantes psicólogos concluyeron exitosamente todas sus tareas sin exceder el número de clics y tiempo máximos. A pesar de ello, reportaron problemas con la lectura de información a causa de un tamaño de fuente, específicamente en el contador del ítem Alertas en el menú y las pestañas correspondientes a la interfaz Resumen Alumno. Debido a esos problemas presentados, reflejaron inconformidad en el cuestionario de satisfacción referente a la visualización de información, con respecto a otros aspectos mostraron un buen nivel de satisfacción.

En la última evaluación se obtuvo un 100 % de tareas completadas exitosamente, un tiempo promedio de tareas de 22.8 segundos, un número de clics promedio de 4.9 y una nota promedio en el cuestionario de satisfacción de 3.947 puntos.

4 Conclusiones

La implementación de las metodologías y herramientas seleccionadas en el desarrollo del prototipo, ayudaron a entender de mejor manera las necesidades de los usuarios. Como prototipo final se obtuvo una aplicación que goza varias funcionalidades permitiendo que el usuario tenga una mejor noción de lo que puede ser el producto final. Además, la evolución que mostró el prototipo fue bastante notoria, pasando de un prototipo básico a uno con interfaces e interacciones muy familiares para los usuarios, además de mejorar en las métricas utilizadas.

Sin embargo, pese a que los resultados obtenidos fueron satisfactorios existen áreas de mejora dentro de las metodologías y técnicas utilizadas. Dentro de estas observaciones se destaca que el número de participantes es menor a lo recomendado, en cuanto a la metodología se puede incorporar principios de diseño más amplios como una evaluación sistemática de la arquitectura de la información y accesibilidad con el propósito de mejorar la navegación y la disponibilidad de información para una mayor diversidad de usuarios como por ejemplo aquellos que padecen algún tipo de discapacidad.

5 Trabajo futuro

El prototipo final tiene aspectos por corregir y áreas en las que se puede mejorar realizando más iteraciones de las etapas "Desarrollo de soluciones de diseño" y "Evaluación del diseño".

Posteriormente, con la aplicación "SelfMind" desplegada, se plantea hacer uso de los datos recolectados por medio de la aplicación para desarrollar un modelo de aprendizaje automático con la finalidad de identificar aquellos estudiantes con más riesgo de padecer tendencias auto-lesivas, además de identificar aquellos factores que más influyen en este fenómeno y posiblemente crear planes de acción dentro de la población de estudiantes.

Agradecimientos. Los autores agradecen a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por la beca otorgada a José Germán González Rodarte. Asimismo, se reconoce el apoyo institucional del Laboratorio de

Investigación en Tecnologías de Experiencia de Usuario (LITUX), por facilitar la infraestructura y el entorno de investigación necesarios para la realización del presente estudio.

Referencias

1. De la Torre IM, Rodríguez JC, Pérez RR, Valdez EA (2009) Ideación suicida en población escolarizada infantil: factores psicológicos asociados. Sistema de Información Científica Redalyc 32(6). Disponible en: <https://www.redalyc.org/articulo.oa?id=58212267007>.
2. Naciones Unidas (2023) El suicidio aumenta en América mientras disminuye en el resto del mundo. Disponible en: <https://news.un.org/es/story/2023/02/1518852>.
3. Instituto Nacional de Estadística y Geografía (INEGI) (2021) Encuesta Nacional de Bienestar Autorreportado. Disponible en: <https://www.inegi.org.mx/programas/enbiare/2021/>.
4. Instituto Nacional de Estadística y Geografía (INEGI) (2023) Día Mundial para la Prevención del Suicidio. Disponible en: https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2023/EAP_Suicidio23.pdf.
5. Revista Colombiana de Psiquiatría (2014) Ideación suicida en estudiantes de medicina: prevalencia y factores asociados. Disponible en: <https://www.elsevier.es/es-revista-revistacolombiana-psiquiatria-379-articulo-ideacion-suicida-estudiantes-medicina-prevalenciaS0034745014000067>.
6. Beautrais A (2020) Stress and suicide in medical students and physicians. N Z Med Stud J 30:11–14. Disponible en: <https://nzmsj.scholasticahq.com/article/12565-stress-and-suicide-in-medical-students-and-physicians>.
7. Jiménez AS, Avelino AA, García CL (2022) Detección temprana del perfil con tendencias suicidas en estudiantes universitarios de la Facultad de Medicina de la Universidad Popular Autónoma del Estado de Puebla asociada a factores de riesgo. Actual Med. Disponible en: <https://digibug.ugr.es/bitstream/handle/10481/78940/am-815-or01.pdf>.
8. Vial S, Boudhraal S, Dumont M (2022) Human-Centered Design Approaches in Digital Mental Health Interventions: Exploratory Mapping Review. JMIR Ment Health 9. Disponible en: <https://mental.jmir.org/2022/6/e35591/PDF>.
9. Wachtler C, Coe A, Davidson S, Fletcher S, Mendoza A, Sterling L, Gunn J (2018) Development of a Mobile Clinical Prediction Tool to Estimate Future Depression Severity and Guide Treatment in Primary Care: User-Centered Design. JMIR Mhealth Uhealth 6(4):e95. Disponible en: <https://mhealth.jmir.org/2018/4/e95>.
10. Galván AS, Reyes SV, Veloz IG, Elizondo PV, González AM, Fierro ML (2022) Development of an Auxiliary Platform (Mentali) for the Primary Screening of Anxiety and Depression in Young Adults. Public Health 19. Disponible en: <https://www.mdpi.com/16604601/19/21/14033>.
11. Norman D (2013) The Design of Everyday Things: Revised and Expanded Edition. Basic Books, New York. Disponible en: <https://kowym.com/wp-content/uploads/2018/08/TheDesign-of-Everyday-Things-Don-Norman.pdf>.
12. International Organization for Standardization (2018) ISO 9241-210: Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts.
13. IEEE (1998) IEEE Std 830-1998: IEEE Recommended Practice for Software Requirements Specifications. IEEE Comput Soc.
14. Cohn M (2004) User Stories Applied: For Agile Software Development. Addison-Wesley.
15. Rubin J, Chisnell D (2008) Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley & Sons.
16. International Organization for Standardization (2018) Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. ISO 9241-11.
17. Shneiderman B (1988) Questionnaire for User Interaction Satisfaction (QUIS). Univ Maryland.
18. Google (2024) Material Design. Disponible en: <https://material.io/design>.



3^{ER} WORKSHOP SOBRE COGNICIÓN E INTERACCIÓN

Aula Virtual Inmersiva de Señalización Deíctica Grupal para Educación Remota

Juan Fernando Florez M.¹, Diego Fernando Valencia M.¹, Elena Muñoz E.¹,
Cesar A. Collazos¹, Julio A. Hurtado¹, Christian Sturm²

¹ Universidad del Cauca, Popayán, Colombia

² Technische Hochschule Ingolstadt, Germany

Resumen. Esta investigación comparó la percepción de inmersión en sistemas de videoconferencia (SVC) convencionales versus inmersivos, durante una actividad de asesoría educativa a distancia. Se diseñó una SVC inmersiva que incorpora flujos de video que varían su posición y señalización deíctica aumentada sobre el material enseñado por parte de los participantes. Se realizaron dos videos de asesoría con un instructor y dos estudiantes que fue observado por 43 estudiantes observadores, quienes evaluaron mediante cuestionarios cuantitativos (escala Likert) y preguntas abiertas. Los resultados cuantitativos (test Mann-Whitney) no mostraron diferencias significativas en la percepción global de inmersión ($p = 0.328$) entre los dos SVC. Sin embargo, el reposicionamiento dinámico de los flujos de video se percibió como un elemento positivo en el SVC inmersivo, mientras que la señalización deíctica aumentada no superó al SVC convencional. El análisis cualitativo reveló ventajas como interacción visual mejorada, pero destacó limitaciones técnicas (latencia, manejo complejo del señalizador) y menor interacción social versus el aula presencial. Se concluye que la flexibilidad espacial en el SVC inmersivo incrementa levemente la percepción de inmersión, pero la implementación técnica actual de señalización no logra un impacto significativo. Futuras investigaciones deberán optimizar la fluidez y naturalidad de las funciones de interacción para potenciar la percepción de inmersión.

Palabras clave: Videoconferencia Inmersiva, Señalización Deíctica, Educación Remota.

1 Introducción

Los sistemas de videoconferencia (SVC) al igual que los sistemas de comunicación (SC) más inmersivos soportados en realidad virtual (VR), aumentada (AR) y extendida (XR) se han usado a pequeña escala en educación remota desde hace décadas [1] [2]. Sin embargo, durante la pandemia COVID19 se hicieron evidentes los problemas de engagement en estudiantes al usar masivamente estas diferentes tecnologías, principalmente los SVC convencionales o comerciales, sin tener en cuenta su diseño simple y los cambios instruccionales que se debían introducir [3][4]. Los resultados de SC inmersivos basados en VR/AR/XR, tampoco fueron mejores, como lo evidencia el estudio comparativo de Bonfert et al., (2023) y la compilación en VR de Mayer et al., (2023).

Los SVC presentan la ventaja de permitir ver los rostros de los participantes durante la reunión [5], y los SC inmersivos tienen la desventaja de la distracción introducida por el mundo virtual [6], al igual que el costo y complejidad hardware introducida por los Head Sets [7]. Bajo el anterior contexto se desarrolló la educación y el trabajo remoto durante pandemia, con participantes experimentando bajos niveles de engagement y fatiga al final del día mientras se usaba masivamente SVC convencionales. Sin embargo, durante este periodo tanto empresas, centros de investigación como investigadores hicieron I+D en SVC [8].

Con respecto a los avances realizados por investigadores en SVC, estos son tanto metodológicos como inmersivos. En la experiencia compilada por Lo et al., (2022) los autores invitan a integrar en los entornos de educación remota elementos metodológicos identificados como exitosos del mundo de videojuegos y streamers como: (a) medir y motivar el desempeño, (b) permitir a los usuarios interactuar directamente con los creadores, (c) captar y mantener el engagement de los usuarios, (d) crear comunidad y (e) curado o selección de los contenidos publicados. Por otra parte, en Katai & Iclanzan, (2023) experimentaron con un formato de videoconferencia de "presencia del instructor en la diapositiva" que permite una mayor expresividad del instructor y una mejor interacción instructor – material enseñado. En [11], se experimentaron con cuatro estilos

de live streaming: a) "cabeza hablante", b) "orador al lado de la diapositiva", c) "orador dentro de la diapositiva" y d) "tablero de luz".

Todos los estilos reportados por los citados autores se usaron en una configuración de educación remota o virtual, donde tanto estudiantes como instructor se encuentran en modalidad remota. En la actualidad hay un alto interés por modelos de educación híbrida, que permitan combinar la virtualidad con la presencialidad para aumentar cobertura, entre otros beneficios [12]. Sin embargo, tanto los modelos virtuales como híbridos conllevan diferentes desafíos por atender que los actuales estilos de video usados en los SVC podrían subsanar. Por ejemplo, en los estilos usados por los citados autores se identifica un potencial incremento en el constructo multidimensional de "presencia" [13] del instructor. Este constructo implica tanto una presencia espacial, una presencia social como una co-presencia [14]. El incremento es particular en las dimensiones de presencia espacial (tele presencia) [15] y presencia social [16] [17] [18] del instructor. Por ejemplo, con respecto a tele presencia en los diferentes estilos se percibe para el "instructor" una sensación de "estar ahí" en el ambiente virtual, como físicamente inmerso. Con respecto a presencia social, es evidente que en la mayor parte de los diferentes estilos de video el instructor comunica mucho mejor sus comportamientos de inmediatez no verbales [19] con la mayor parte de su cuerpo facilitando la señalización deíctica sobre el material enseñado. Característica esta que es mínima con el estilo "cabeza hablante", típico de las SVC convencionales, cuando se comparte pantalla con los estudiantes y el video del instructor se reduce aún más en tamaño. Pero, con respecto a la co-presencia [20] [21] en todos los estilos de video de los autores citados esta es mínima. Esto último debido a la naturaleza asimétrica del medio de comunicación [22] usado en todos los estilos, ya que solamente hay percepción de inmersión y capacidad de señalización deíctica en el entorno virtual en el camino hacia el instructor y no hacia el resto de los participantes en los SVC inmersivos reportados por los citados autores.

El presente trabajo aborda el desafío de la inmersión y señalización deíctica aumentada no solo del instructor sino también del resto de los participantes en una actividad de asesoría en educación remota, como medio para aumentar los niveles de engagement en reuniones virtuales mediadas por SVC. El artículo presenta en la sección 2 el método y materiales relacionados. En las secciones 3 y 4 se presentan los resultados de percepción en diseño de inmersión y una discusión de los resultados. Finalmente, en la sección 5 se muestran las conclusiones y trabajos futuros.

2 Método y materiales

Se realizó una actividad de asesoría entre dos estudiantes y un instructor replicada en dos sistemas de videoconferencia: inmersiva y no inmersiva. Se invitaron a dos grupos de estudiantes observadores para responder un cuestionario. Se utiliza un enfoque de método mixto, donde se recopilan datos cuantitativos y cualitativos para realizar un análisis inferencial de la percepción de los estudiantes observadores sobre la experiencia de inmersión realizada en cada SVC y sus recomendaciones al respecto.

2.1 Participantes, observadores y SVC

Los participantes de la actividad de asesoría son un instructor y dos estudiantes de un curso de instrumentación industrial de un programa STEM de educación superior. Estos realizan la actividad en una reunión virtual mediada por una SVC. Esta consiste en solucionar dudas sobre cómo "leer" o interpretar diagramas de proceso e instrumentación (P&ID). Se realizaron dos réplicas de una misma actividad de asesoría, mediadas por una tecnología de SVC diferente: la primera es una SVC inmersiva que permite señalización deíctica aumentada sobre el material enseñado y la segunda es una SVC convencional. En ambos casos se grabó la actividad de asesoría durante 5 minutos.

Los observadores son estudiantes invitados del mismo programa STEM de diferentes semestres, lo que permite una muestra heterogénea de estudiantes de educación superior, con experiencia previa con plataformas de videoconferencia. Estos estudiantes observadores fueron divididos en dos grupos para observar un video de cinco minutos de una de las asesorías y a continuación responder voluntariamente un cuestionario.

La SVC inmersiva esta soportada en la herramienta comercial Zoom, habilitando su vista inmersiva y procesando previamente el flujo de video de cada participante para permitir una señalización deíctica aumentada más allá de los límites del flujo de video sobre el material enseñado. La SVC no inmersiva es la herramienta comercial *Google Meets*. En ambos casos se usa la opción de compartir pantalla para presentar un P&ID, este constituye el material enseñado sobre los cuales se realiza la actividad de asesoría (ver Fig.1).

Las Figuras 1.c y d ilustran las características diferenciales de la SVC inmersiva: a) flujos de video de los participantes flotando libremente sobre el material enseñado, b) coherencia espacial entre los flujos de video y el material enseñado y c) señalización deíctica aumentada más allá del marco original del flujo de video del participante. El procesamiento del flujo de video adicional realizado en esta investigación al SVC inmersivo permite que cada participante adicional a las características b) y c) ya descritas pueda: 1) Desplazarse a posiciones fijas sobre el material enseñado, 2) escribir sobre el material enseñado, 3) resaltar sobre el material enseñado, 4) invertir su flujo de video y 5) modificar la extensión de su señalizador deíctico. Las características b) y c) y las cinco funcionalidades señaladas no se cumplen en el SVC convencional como se ilustra en las Figuras 1.a y b.

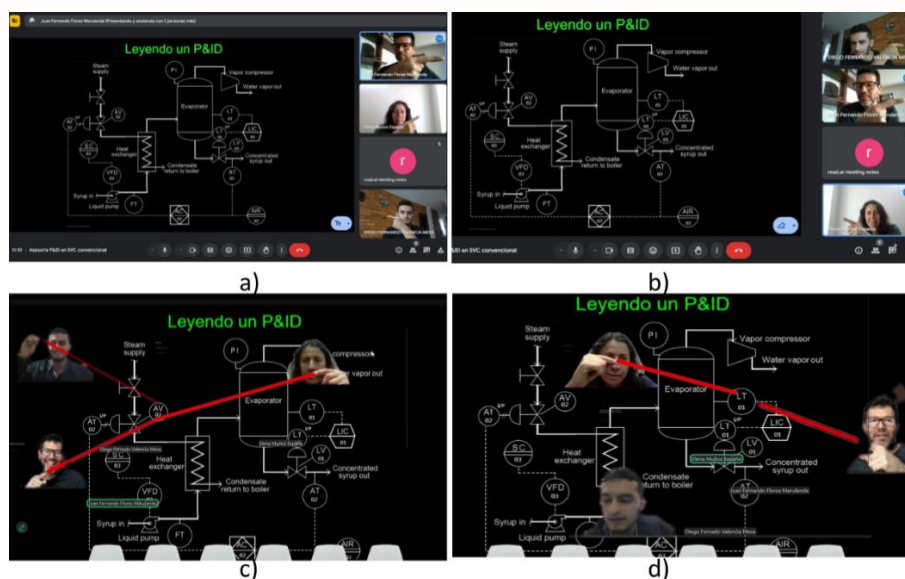


Fig. 1. Tecnologías de SVC usadas durante la actividad de asesoría. a) y b) SVC convencional sin coherencia espacial y vistas diferentes para cada participante. c) y d) SVC inmersiva, coherencia espacial y señalización deíctica aumentada.

2.2 Procedimiento y cuestionario

Se diseñó un protocolo de interacción entre los participantes durante la asesoría para aprovechar las características implementadas en la SVC inmersiva. En este el participante host (instructor) ubica libremente los flujos de video del resto de participantes estudiantes sobre zonas específicas del material enseñado y usa señalización deíctica para explicar (ver Fig. 1 c). Una vez ubicado en la zona de interés el estudiante participante también puede usar señalización deíctica para ubicar su duda o explicar lo que entiende al respecto (ver Fig. 1 d). En este proceso el host puede ajustar posición y tamaño del flujo de video del estudiante participante para facilitar la señalización deíctica del mismo sobre la zona de interés (ver Fig. 1 d y c). De esta forma los participantes pueden usar tanto comportamientos verbales como no verbales en sus procesos de comunicación e interacción durante la actividad de asesoría remota.

Los procesos de comunicación e interacción de la actividad de asesoría realizada en la SVC convencional se llevaron a cabo acorde a las limitaciones que este tipo de tecnologías implica. En ambos casos se grabaron videos de 5 minutos de la actividad de asesoría realizada en cada SVC. Estos videos se subieron a *YouTube* y

un enlace fue incrustado en un formulario. Cada enlace del formulario fue enviado vía email a dos grupos de estudiantes STEM. Cada destinatario podía decidir individualmente si participar o no en la encuesta, lo que resultó en una tasa de respuesta promedio del 45,55 %. No se realizó presión social en los estudiantes para diligenciar la encuesta y así evitar consecuencias negativas por responder sin sinceridad. El cuestionario comenzó con un formulario de consentimiento. Luego se preguntó si ya tenía experiencia previa con plataformas de videoconferencia, obteniendo que el 95,34% de los participantes si la tenía.

Para recopilar los datos cuantitativos se usó un instrumento validado para evaluar la aceptación de un entorno virtual de aprendizaje basado en Moodle. El instrumento fue creado por I. Ruiz & S. Romero, (2008) y validado por Santana-Mancilla et al., (2019) y consta de cinco factores: comunicación, diseño, usabilidad, aspectos generales y fiabilidad, con 15 ítems, tres por cada factor. Para adecuar el instrumento a la evaluación del diseño de inmersión en plataformas de videoconferencia se usó el factor diseño (de inmersión) modificando y/o ajustando la redacción en cada uno de los tres ítems hacia un SVC (ver Tabla 1). Para responder al factor se usaron tres ítems en una escala de Likert de 7 niveles (1-Totalmente en desacuerdo, 2- Bastante en desacuerdo, 3-En desacuerdo, 4-Neutral, 5-De acuerdo, 6-Bastante de acuerdo, 7- Totalmente de acuerdo) (ver Tabla 1).

Table 2. Factor y tres ítems utilizados para evaluar la SVC (inmersiva y no immersiva).

Diseño (inmersión)	Observar los flujos de video del profesor y de los estudiantes en la SVC fue útil para el desarrollo de la asesoría.
	Los cambios de posición y tamaño de los flujos de video de los participantes es adecuado para percibir que se sienten inmersos.
	La señalización deíctica usada en la SVC fue útil para realizar la asesoría.

Adicionalmente, el cuestionario contiene preguntas abiertas sobre las apreciaciones de ventajas y desventajas de la SVC immersiva diseñada, en comparación con un salón de clase o una SVC convencional.

2.3 Análisis de datos

Los cuestionarios, adicional a la información demográfica, recogieron en tres preguntas datos cuantitativos (ver Tabla 1) y en cuatro preguntas información cualitativa (respuestas a preguntas abiertas). A los datos cuantitativos se les hace un análisis inferencial mediante un diseño experimental comparativo entre un control (SVC convencional – SVC C) y un tratamiento (SVC immersivo – SVC A) en grupos independientes de estudiantes observadores. En este se analiza la diferencia en la percepción de los observadores del diseño de inmersión de las dos tecnologías SVC. Los datos cualitativos de los cuestionarios se analizaron usando análisis temático mediante la identificación de similitudes y la creación inductiva de nubes de palabras. Se recopilaron 43 (19 SVC C y 24 SVC A) respuestas con datos cualitativos además de los tres ítems del cuestionario. El presente análisis plantea dar respuesta a la siguiente pregunta:

Q. ¿Cuáles son los efectos en la percepción de inmersión de observar los flujos de video de los participantes flotando y usar señalización deíctica aumentada sobre el material enseñado cuando se usa un sistema de videoconferencia?

El diseño experimental implicó hacer uso de un preprocesamiento en Python a los flujos de video de los participantes para generar un flujo de video aumentado, emplear OBS como generador de una cámara virtual, utilizar los SVC comerciales *Zoom* y *Google Meets* y finalmente *Google Forms* para generar el formulario de encuesta. Con Python se realizó el tratamiento estadístico de los datos cuantitativos e igualmente un análisis temático y gráfico de las respuestas a las preguntas abiertas.

3 Resultados y discusión

Con el fin de dar respuesta a la pregunta de investigación planteada se analiza la distribución de los tres ítems del factor Diseño de Inmersión, para posteriormente realizar una comparación estadística del SVC inmersivo en contraste con el SVC convencional.

3.1 Análisis de distribución y preprocesamiento de los datos

Se llevó a cabo una prueba de normalidad Shapiro-Wilk de cada uno de los tres ítems del factor Diseño de Inmersión en los grupos SVC C y SVC A, resultando que ninguno de los seis ítems presenta distribución normal. De acuerdo con lo anterior se aplicó un método de sustitución de atípicos por rango Inter cuartil, cada valor fuera del rango Inter cuartil fue reemplazado por la mediana, ver resultado en la Figura 2.

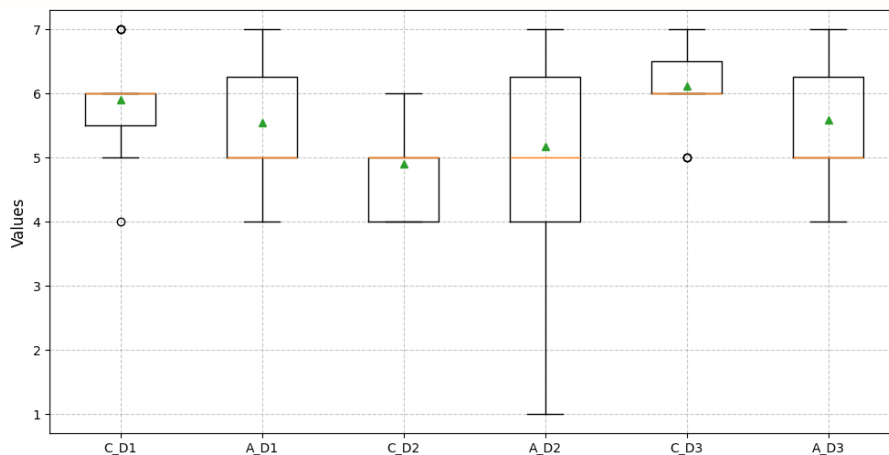


Fig. 2. Diagramas de cajas y bigotes comparativo de cada uno de los tres ítems (D1 a D3) de los datos SVC C y SVC A.

3.2 Comparación SVC inmersivo y convencional

Se llevó a cabo una prueba de normalidad Shapiro-Wilk a los datos del factor diseño de inmersión de SVC C y SVC A, hallando que en ambos casos siguen una distribución no normal. Por lo que se aplica una prueba Mann-Whitney para determinar si existen diferencias significativas entre los factores de diseño de inmersión percibidos entre SVC C y SVC A. El estadístico Mann-Whitney es 2250.5 y el p-value es 0.3289, por lo que no se presentan diferencias significativas.

3.3 Análisis temático y nubes de palabras de preguntas abiertas

Con el propósito de ahondar en aspectos técnico o de otra índole implicados en la percepción de inmersión de los estudiantes observadores se diseñaron cuatro preguntas abiertas. De interés particular, son las cuatro preguntas abiertas que el grupo SVC A respondió, las cuales fueron procesadas con análisis temático y nubes de palabras:

- P1. ¿Encuentra ventaja de la SVC A con respecto al aula de clase real, mencionar?
- P2. ¿Encuentra ventaja de la SVC A con respecto a un SVC convencional, mencionar?
- P3. ¿Encuentra desventaja de la SVC A con respecto al aula de clase real, mencionar?
- P4. ¿Encuentra desventaja de la SVC A con respecto a un SVC convencional, mencionar?

Para la P1 los temas resaltados como ventaja en su orden son: ninguna, puede que facilite la comunicación y la asesoría, menos desplazamiento y comodidad de encontrarse. Con respecto a P2 los temas son: similar, inmersiva y didáctica, facilidad de compartir y fácil señalización. Para P3 los temas resaltados como desventaja son: prestar atención, ninguna, menor interacción y problemas de conexión. Finalmente, para la P4 los temas principales son: de pronto, la intermitencia del internet, limitaciones en las funcionalidades y ninguna. Las nubes de palabras asociadas a los temas principales de las cuatro preguntas abiertas en SVC A se observan en la Figura 3.

El análisis del diagrama de cajas y bigotes de cada uno de los ítems (D1 a D3) entre SVC C y SVC A, señala que tanto observar los flujos de video del instructor-estudiantes (D1) como usar la señalización deíctica aumentada (D3) no fue superior en utilidad con respecto al SVC convencional, solamente los cambios de posición y tamaño de los flujos de video de los participantes (D2) fue considerado por los observadores que permite una mayor sensación de inmersión con respecto al SVC convencional, pero no lo suficiente para ser una diferencia significativa. Aún más, tanto las medias como las medianas de D1 y D3 en SVC A están por debajo de las D1 y D3 de SVC C. Esto permite explicar que no haya diferencias significativas en el factor diseño de inmersión de SVC A con respecto a SVC C.



Fig. 3. Nubes de palabras asociadas a los temas principales de las cuatro preguntas abiertas en SVC A.

Finalmente, ello permite responder a la pregunta de investigación planteada señalando que solo hay un leve efecto positivo en la percepción de inmersión en la SVC A cuando se cambian de posición y tamaño los flujos de video de los participantes, pero no cuando se usa señalización déctica aumentada. A partir del análisis a los temas principales identificados en las respuestas a las cuatro preguntas abiertas se infiere que aún hay aspectos del diseño de inmersión en el preprocesamiento de los flujos de video de los participantes que se deben corregir, para que la interacción con el señalizador sea más natural y particularmente para evitar retrasos o intermitencia en el flujo de video del SVC inmersivo.

4 Conclusiones y trabajo futuro

Se diseñó un sistema de videoconferencia aumentada sobre una plataforma comercial que incorpora en los flujos de video de los participantes una misma coherencia espacial, esto permite entre todos los participantes la

capacidad de señalización deíctica aumentada sobre el material explicado en diapositivas, y la libertad de posicionar los flujos de video en cualquier sitio sobre el material enseñado.

El análisis estadístico realizado entre el actual diseño de la SVC inmersiva y la SVC convencional señala al elemento de diseño de inmersión “cambiar la posición y tamaño de los flujos de video de los participantes” como un ítem positivo en la percepción de inmersión entre los participantes observadores. Pero no así, el uso del elemento de diseño “señalización deíctica aumentada”. El análisis cualitativo temático de las ventajas y desventajas del SVC inmersivo con respecto al aula presencial señala comodidad, menos transporte, pero igualmente menor interacción social. El análisis temático de las ventajas y desventajas del SVC inmersivo con respecto al SVC convencional realizado en las cuatro preguntas abiertas, resalta la interacción visual mejorada pero también permite señalar que la complejidad en el manejo del señalizador aumentado por parte de los participantes y principalmente el retraso o intermitencia variable en los flujos de video son aspectos de diseño que se deben atender en futuros trabajos.

Referencias

1. C. A. Warden, J. O. Stanworth, J. B. Ren, and A. R. Warden, “Synchronous learning best practices: An action research study,” *Comput Educ*, vol. 63, pp. 197–207, 2013, doi: 10.1016/j.compedu.2012.11.010.
2. P. Pérez, E. González-Sosa, J. Gutiérrez, and N. García, “Emerging Immersive Communication Systems: Overview, Taxonomy, and Good Practices for QoE Assessment,” in *Frontiers in Signal Processing*, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:248721669>.
3. P. R. Lowenthal, J. Borup, R. E. West, and L. Archambault, “Thinking Beyond Zoom: Using Asynchronous Video to Maintain Connection and Engagement During the COVID-19 Pandemic,” 2020. [Online]. Available: <https://www.marcopolo.me>.
4. G. Orhan and Ö. Beyhan, “Teachers’ perceptions and teaching experiences on distance education through syn-chronous video conferencing during covid-19 pandemic,” *Social Sciences and Education Research Review*, vol. 1, no. 7, pp. 8–44, 2020, [Online]. Available: www.sserr.ro.
5. M. Bonfert *et al.*, “Seeing the faces is so important—Experiences from online team meetings on commercial virtual reality platforms,” *Front Virtual Real*, vol. 3, 2023, doi: 10.3389/frvir.2022.945791.
6. R. E. Mayer, G. Makransky, and J. Parong, “The Promise and Pitfalls of Learning in Immersive Virtual Reality,” *Int J Hum Comput Interact*, vol. 39, no. 11, pp. 2229–2238, 2023, doi: 10.1080/10447318.2022.2108563.
7. Philippe Fuchs, Judith Guez, Olivier Hugues, Jean-Francosi Jégo, Andras Kemeny, and Daniel Mestre, *VIRTUAL REALITY HEADSETS - A theoretical and Pragmatic Approach*. London, UK: CRC Press/Balkema, 2017.
8. M. A. Camilleri and A. C. Camilleri, “Remote learning via video conferencing technologies: Implications for research and practice,” *Technol Soc*, vol. 68, Feb. 2022, doi: 10.1016/j.techsoc.2022.101881.
9. A. W. Lo, B. Stevens, and S. P. Willems, “World of EdCraft: Challenges and Opportunities in Synchronous Online Teaching,” *Harv Data Sci Rev*, Apr. 2022, doi: 10.1162/99608f92.73a1c910.
10. Z. Katai and D. Iclanzan, “Impact of instructor on-slide presence in synchronous e-learning,” *Educ Inf Technol (Dordr)*, vol. 28, no. 3, pp. 3089–3115, 2023, doi: 10.1007/s10639-022-11306-y.
11. J. F. Florez Marulanda, “Student Satisfaction Pilot Experience with Synchronous Classroom Live Streaming Styles during the COVID-19 Pandemic,” *Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 17, no. 3, pp. 301–306, 2022, doi: 10.1109/RITA.2022.3191285.
12. M. W. Bülow, “Designing Synchronous Hybrid Learning Spaces: Challenges and Opportunities,” 2022, pp. 135–163. doi: 10.1007/978-3-030-88520-5_9.
13. M. Lombard, T. B. Ditton, and L. Weinstein, “Measuring Presence: The Temple Presence Inventory,” 2009. [Online]. Available: <https://www.researchgate.net/publication/228450541>.
14. S. T. Bulu, “Place presence, social presence, co-presence, and satisfaction in virtual worlds,” *Comput Educ*, vol. 58, no. 1, pp. 154–161, Jan. 2012, doi: 10.1016/j.compedu.2011.08.024.
15. T. W. Schubert, “A new conception of spatial presence: Once again, with feeling,” *Communication Theory*, vol. 19, no. 2, pp. 161–187, 2009, doi: 10.1111/j.1468-2885.2009.01340.x.
16. Jhon Short, Ederyn Williams, and Bruce Christie, *The social psychology of telecommunications*. London, UK: Wiley, 1976.
17. C. S. Oh, J. N. Bailenson, and G. F. Welch, “A systematic review of social presence: Definition, antecedents, and implications,” 2018, *Frontiers Media S.A.* doi: 10.3389/frobt.2018.00114.

18. Juan F. Florez M., Julio A. Hurtado, and Cesar A. Collazos, "Social Presence and User Experience: The Influence of the Immersive Virtual Classroom in Synchronous Distance Learning," in *IX Iberoamerican Conference on Human Computer Interaction*, Buenos Aires, 2023, pp. 1–15.
19. M. Schutt, B. S. Allen, and M. A. Laumakis, "The effects of instructor immediacy behaviours in online learning environments," *The Quarterly Review of Distance Education*, vol. 10, no. 2, pp. 135–148, 2009.
20. E. Goffman, *Behavior in public places*. New York, 1963.
21. R. Schroeder *et al.*, "Collaborating in networked immersive spaces: as good as being there together?," 2001.
22. M. Billinghurst, S. Bee, J. Bowskill, and H. Kato, "Asymmetries in Collaborative Wearable Interfaces," 1999, pp. 133–140.
23. I. Ruiz and S. Romero, "Moodle: una herramienta eficaz aplicada a la enseñanza de prácticas, en el área de electrónica y arquitectura de computadores," VIII Congreso de Tecnologías Aplicadas a la Enseñanza de Electrónica, 2008, p. 165.
24. P. C. Santana-Mancilla, O. A. Montesinos-López, M. A. García-Ruiz, J. J. Contreras-Castillo, and L. S. Gaytan-Lugo, "Validación de un instrumento para medir la aceptación tecnológica de un entorno virtual de aprendizaje," *Acta Univ*, vol. 29, 2019, [Online]. Available: <http://doi.org/10.15174.au.2019.1796>



2º WORKSHOP SOBRE ACCESIBILIDAD E INCLUSIÓN

Day a Day Communicator: Aplicación Web Destinada a Mejorar la Comunicación Asistida de Personas con Trastorno del Espectro Autista

Cristina Collado¹, Susana Bautista¹

¹Escuela Politécnica Superior, Universidad Francisco de Vitoria, Madrid (España)
ccolladop00@gmail.com, susana.bautista@ufv.es

Resumen. Este proyecto se centra en el desarrollo de "Day a Day Communicator²", una aplicación web destinada a mejorar la comunicación asistida de personas con Trastorno del Espectro Autista (TEA). La creación de esta aplicación se ha basado en un diseño centrado en el usuario, con la participación de los expertos del Colegio CEPRI³ a lo largo de todo el proceso. Desde el inicio, se realizaron reuniones para capturar los requisitos específicos y comprender las necesidades de los usuarios finales, que son los alumnos del Colegio CEPRI. Durante el desarrollo, se elaboraron varios prototipos que fueron evaluados y refinados según las observaciones de los especialistas. El diseño se centró en la personalización y accesibilidad, incorporando elementos como un lector grande, navegación intuitiva y opciones de personalización con imágenes individuales. Además, se optimizó para ser web responsive, garantizando una experiencia fluida en diversos dispositivos, como ordenadores, tablets y algunos móviles. La aplicación fue desarrollada utilizando tecnologías como HTML, CSS, JavaScript y React. Al finalizar el desarrollo, se realizó una evaluación final con los especialistas y los usuarios finales, quienes proporcionaron un feedback positivo sobre la utilidad y facilidad de uso de la aplicación. "Day a Day Communicator" ha sido ampliamente valorada por su capacidad para ofrecer una solución personalizada y efectiva para la comunicación diaria de personas con TEA.

Keywords: Comunicación Asistida, TEA, Aplicación Web, Personalización, Accesibilidad, Diseño Centrado en el Usuario.

1 Introducción

En los últimos tiempos se ha dedicado una considerable atención a la integración de las TIC en el ámbito educativo. Se han observado los beneficios que pueden abarcar en la enseñanza por parte del profesorado, especialmente en el trabajo con alumnado que presenta TEA. Según datos del Ministerio de Educación de 2022 [1], los estudiantes con autismo representan el 28,1% del total de alumnos con necesidades educativas especiales. Esto resalta la importancia de utilizar herramientas flexibles y personalizadas en el proceso de enseñanza, adaptadas al ritmo y las características individuales de cada alumno.

El Colegio C.E.E. CEPRI, tiene como fin social promover la investigación y el estudio de aquellos factores, incidentes, causantes o relacionados con el autismo, Trastorno del Espectro Autista u otros Trastornos Generalizados del Desarrollo, así como su rehabilitación o prevención. En el centro hay 56 alumnos con necesidades de apoyo, de los cuales el 75% se comunican mediante lenguaje no verbal mientras que el resto presentan dificultades de comunicación. Ante esta realidad, resulta imperativo proporcionar un sistema de comunicación alternativo o aumentativo que satisfaga las necesidades de estos estudiantes.

Un comunicador es una herramienta diseñada para facilitar la comunicación de personas con dificultades para expresarse verbalmente. Los comunicadores tradicionales, como los basados en pictogramas impresos, están siendo reemplazados por sistemas digitales más avanzados que permiten una mayor personalización. Estos dispositivos permiten a los usuarios seleccionar pictogramas, palabras o frases mediante pantallas táctiles y reproducir el mensaje de forma audible, adaptándose a las necesidades específicas de cada usuario.

¹ <https://www.ufv.es/>

² <https://dayadaycommunicator.onrender.com/>

³ <http://colegiocepri.com/>

Con el propósito de abordar esta necesidad, se ha emprendido el desarrollo de un sistema de comunicación, para personas de cualquier edad y género, que busca mejorar las soluciones existentes, adaptándolas a las necesidades específicas identificadas por los expertos en el campo. Para ello, se ha tomado como referencia el framework Asterics Grid⁴, el cual, si bien es una base sólida, carece de ciertas funciones requeridas para cubrir las necesidades particulares de los alumnos del centro. Se tiene previsto que en trabajos futuros se establezca una conexión entre ambas soluciones, integrando así lo mejor de cada una para ofrecer un sistema de comunicación más robusto y efectivo para los estudiantes con necesidades especiales.

2 Estado del arte

Ponemos el foco en la brecha digital al colectivo con el que estamos trabajando, presentamos la discapacidad intelectual y algunas aplicaciones relacionadas con la propuesta que se presenta.

El proyecto "Day a Day Communicator" nace como una respuesta tecnológica para mejorar la comunicación de los niños con TEA buscando una solución accesible, eficaz y personalizada que se adapte a las necesidades de este colectivo, en particular a los alumnos del Colegio CEPRI. Su desarrollo ha seguido un enfoque iterativo y un diseño centrado en el usuario (DCU), con el objetivo de crear una herramienta sencilla pero potente para facilitar la interacción y expresión de los niños en su vida cotidiana. Esta metodología es especialmente eficaz para proyectos que buscan mejorar la experiencia del usuario, ya que permite ajustar el producto en función de la retroalimentación directa y continua. Además, el DCU permite minimizar riesgos al alinearse estrechamente con las expectativas y necesidades del usuario, garantizando así que el producto final sea tanto accesible como funcional.

2.1 Trastorno del espectro autista

El Trastorno del Espectro Autista (TEA) es una discapacidad del desarrollo [1] que afecta diversas áreas como la comunicación, el lenguaje y la interacción social. Las personas con TEA experimentan dificultades para adaptarse a cambios en su rutina, presentan sensibilidades sensoriales atípicas y exhiben comportamientos repetitivos. Sin embargo, su manera de aprender, interactuar y comunicarse puede ser única, lo que hace que el TEA se manifieste de manera diversa en cada individuo.

El TEA tiene una base genética significativa [2], lo que sugiere que las variantes genéticas juegan un papel importante en el desarrollo del trastorno. Estudios de gemelos han mostrado una alta concordancia en gemelos monocigóticos (90%) frente a gemelos dicigóticos (20%), lo que refuerza la hipótesis genética. Sin embargo, factores ambientales también contribuyen al desarrollo del TEA. Estos factores incluyen condiciones prenatales, como la salud materna, la edad de los padres, la exposición a infecciones o el uso de medicamentos durante el embarazo, que pueden interactuar con la predisposición genética.

2.2 Dificultades en la comunicación

Las personas con TEA enfrentan diversas dificultades [8]. Una de las principales es la comunicación, ya que los individuos con TEA pueden presentar problemas tanto en el lenguaje verbal como no verbal, lo que dificulta la interacción social. Además, muestran comportamientos repetitivos y restringidos que pueden interferir con las actividades diarias y la capacidad de adaptarse a cambios en su entorno. También es común que presenten sensibilidades sensoriales atípicas, como hipersensibilidad o hiposensibilidad a estímulos como la luz, el sonido o el tacto. Por último, muchas personas con TEA tienen dificultades para planificar y organizar tareas complejas, lo que puede dificultar aún más su capacidad para adaptarse a cambios inesperados.

⁴ <https://grid.asterics.eu/#main>

2.3 Sistemas alternativos y aumentativos de comunicación (SAACs)

Para superar estas barreras, uno de los enfoques más efectivos es el uso de Sistemas Alternativos y Aumentativos de Comunicación (SAAC) [4]. Estos sistemas están diseñados para ayudar a las personas con dificultades en el habla a expresarse de manera efectiva. Utilizan formas alternativas de comunicación, como símbolos, pictogramas, tarjetas o gestos [9]. Los SAACs son herramientas muy útiles, ya que no solo facilitan la expresión, sino que también promueven la independencia de las personas con TEA al proporcionarles una manera estructurada y visualmente clara de comunicarse.

Dentro de los SAAC, se pueden encontrar varias opciones. El Sistema Pictográfico de Comunicación (SPC)⁵ utiliza imágenes o pictogramas que las personas pueden señalar para comunicarse. Este sistema es útil para aquellas personas con dificultades cognitivas o de lenguaje, ya que les ofrece una representación visual de la información. Por otro lado, el Sistema de Ideogramas y Pictogramas de Comunicación (PIC)⁴ organiza los pictogramas en categorías más amplias, lo que permite una estructura jerárquica y facilita la comprensión de conceptos más complejos.

2.4 Aplicaciones similares

Las herramientas y aplicaciones son ejemplos de cómo la tecnología y los sistemas alternativos de comunicación pueden mejorar significativamente la calidad de vida de las personas con TEA. Estos avances no solo contribuyen a superar las barreras de comunicación, sino que también promueven la inclusión social, la autonomía y la participación activa de las personas con TEA en la sociedad, permitiéndoles expresarse y conectarse de manera más efectiva con el mundo que les rodea.

Existen plataformas digitales como Arasaac [3], que destaca por ser una herramienta gratuita que proporciona una extensa colección de pictogramas y recursos gráficos para facilitar la comunicación y la accesibilidad cognitiva. Su uso es ampliamente extendido en contextos educativos y terapéuticos, promoviendo la inclusión de personas con necesidades especiales de comunicación. AsTeRICS Grid [11], Let Me Talk⁶, Niki Talk⁷, Symbo Talk⁸ son aplicaciones que utilizan pictogramas y símbolos visuales para permitir a los usuarios construir frases y expresar sus necesidades de manera efectiva.

Teniendo en cuenta estos ejemplos de aplicaciones similares ya diseñadas en el marco del estudio y en colaboración con los expertos del Colegio CEPRI, se identificaron las limitaciones y desafíos de los comunicadores tradicionales que no lograban una integración adecuada con las necesidades individuales de los usuarios. Esto impulsó la creación de un prototipo que, con una interfaz amigable e intuitiva, que mejoraba la experiencia de comunicación a través de pictogramas categorizados. Estos pictogramas se agrupan en tres categorías principales: lugares, verbos y acciones, permitiendo que los niños pudieran expresar ideas o hacer peticiones con mayor facilidad. Además, el prototipo incorporaba una funcionalidad de lectura de frases, lo que proporcionaba un medio auditivo de apoyo en el proceso comunicativo. También se incorporaron opciones de personalización, como la posibilidad de ajustar la interfaz según las preferencias de cada usuario. Este enfoque buscaba garantizar que los usuarios pudieran modificar la herramienta para adaptarla a sus propias necesidades, una característica fundamental en el desarrollo de tecnología para personas con TEA.

En conclusión, Day a Day Communicator, no se limita a ofrecer un repositorio de pictogramas, sino que organiza la comunicación en categorías adaptadas a contextos reales. Su diseño centrado en el usuario y adaptado a perfiles cognitivos y sensoriales específicos proporciona una experiencia más accesible, estructurada y significativa para las personas con TEA, mejorando su autonomía y capacidad de expresión en el día a día.

⁵ <https://www.mheducation.es/bcv/guide/capitulo/8448183797.pdf>

⁶ https://aulaabierta.arasaac.org/wp-content/uploads/letmetalk/Manual_LetMeTalk.pdf

⁷ <http://www.nikitalk.com/>

⁸ <https://www.symbotalk.com/spanish/spanish.html>

3 Implementación del day and day communication

En términos de implementación tecnológica, la aplicación fue desarrollada utilizando HTML, CSS y React para la creación del frontend. React⁹, al ser una biblioteca popular para el desarrollo de interfaces de usuario, permitió un enfoque modular y flexible, adaptándose a las necesidades de la plataforma y optimizando la experiencia tanto en dispositivos móviles como en escritorios. Para la interactividad y la gestión del estado, se utilizó JavaScript, asegurando que la aplicación fuera dinámica y respondiera en tiempo real a las interacciones de los usuarios. Además, el diseño se implementó siguiendo los principios del diseño responsive, garantizando que la aplicación fuera completamente funcional en diferentes tamaños de pantalla, permitiendo su uso en tablets, móviles y ordenadores de escritorio.

Para el despliegue, se utilizó Git¹⁰ y Render¹¹. Git permitió una gestión eficiente del código fuente, mientras que Render, al ser una plataforma en la nube, permitió una integración continua y actualizaciones frecuentes del sistema. La nube garantiza la accesibilidad desde cualquier lugar, y su infraestructura robusta asegura la estabilidad del servicio para los usuarios finales. Esto facilita que los educadores, padres o cuidadores puedan acceder a la herramienta en cualquier momento y desde cualquier dispositivo conectado a Internet.

El diseño final del comunicador incluye una mejora en la disposición de los elementos en pantalla, con un lector de frases ampliado para facilitar la lectura y una barra deslizante que permite una navegación más fluida. También se facilitó la importación de imágenes personalizadas por parte de los usuarios, lo que amplió la versatilidad de la herramienta.

4 Evaluación

Con el fin de mejorar la herramienta y hacerla lo más eficiente posible, se llevó a cabo un proceso de pruebas constantes con los expertos. Las pruebas iterativas permitieron implementar cambios en tiempo real, asegurando que la herramienta evolucionara según las necesidades detectadas. El prototipo inicial fue determinado por la participación de educadores y expertos en el área del TEA del Colegio CEPRI, quienes proporcionaron valiosa retroalimentación para optimizar la herramienta, tanto a nivel de diseño como de funcionalidad, a lo largo del proceso hasta conseguir un prototipo final con el que se ha evaluado la aplicación. Entre las evaluadoras se encontraban la psicóloga, la pedagoga, la logopeda y docentes con experiencia en comunicación aumentativa y alternativa (CAA).

Se realizaron las evaluaciones mediante dos tipos de cuestionarios. Los cuestionarios de funcionalidad¹² permitieron analizar aspectos específicos de la usabilidad y personalización, destacando que la facilidad de uso y la personalización mediante imágenes fueron altamente valoradas. Por otro lado, los cuestionarios SUS¹³ (System Usability Scale[6]) proporcionaron un análisis cuantitativo de la experiencia de usuario. Los expertos coinciden en que la funcionalidad de personalización de la aplicación es altamente útil, destacando su facilidad para adaptar la comunicación, la posibilidad de añadir imágenes individuales y su uso intuitivo. La herramienta obtuvo un 100% de aceptación en estas funcionalidades y un 75% de valoración positiva en cuanto a facilidad de uso por parte de los usuarios finales. Se reconoce que la efectividad de un comunicador destinado a personas con TEA reside precisamente en su capacidad de personalización e individualización. Este enfoque permite adaptar la herramienta a las características únicas de cada usuario, lo que representa una base sólida para futuras evaluaciones centradas en el progreso comunicativo.

⁹ <https://es.react.dev/>

¹⁰ <https://github.com/>

¹¹ <https://render.com/>

¹² <https://forms.office.com/r/ys0Zs2Ekiu>

¹³ <https://forms.office.com/r/kg9vZGYuPR>

La evolución del proyecto se basará en un enfoque iterativo, donde cada actualización responderá directamente a las necesidades detectadas en la evaluación. Uno de los principales desafíos detectados es la necesidad de un sistema de administración que permita gestionar comunicadores de forma permanente. Para abordar esta cuestión, se plantea la implementación de una base de datos y una API que permitan a los usuarios almacenar y recuperar sus configuraciones, asegurando una experiencia más fluida y adaptada a sus necesidades. Es por ello, que la integración de perfiles de usuario ha sido señalada como una necesidad clave.

Estas mejoras no solo optimizarán la experiencia de los usuarios actuales, sino que también permitirán que la herramienta siga creciendo y adaptándose a nuevos contextos y requerimientos.

5 Conclusiones y trabajo futuro

El proyecto "Day a Day Communicator" ha logrado cumplir sus objetivos iniciales al ofrecer una herramienta accesible para mejorar la comunicación de personas con TEA.

El respeto a la diversidad de cada persona con TEA también es crucial. Cada individuo tiene sus propias necesidades, habilidades y preferencias, y es importante proporcionar un apoyo personalizado y centrado en la persona. Esto implica reconocer y valorar la singularidad de cada persona, así como promover su inclusión y participación en la sociedad.

El éxito de este enfoque, radica en reconocer que, al adaptar la tecnología a las necesidades individuales, se potencia no solo la comunicación, sino también la autonomía y la calidad de vida de las personas con TEA.

La colaboración con los expertos del Colegio CEPRI fue fundamental para que el proyecto siga evolucionando con el objetivo de ampliar su alcance, incorporando nuevas funcionalidades que mejoren aún más la accesibilidad y favorezcan la integración social de los usuarios. En consecuencia, en el futuro, se prevé implementar un sistema de administración que permita crear y gestionar comunicadores permanentes, así como desarrollar una base de datos y API que ofrezcan una gestión más eficiente y segura de los datos de los usuarios. Además, se considerará la creación de perfiles de usuario que permitan a cada persona guardar sus preferencias y personalizar opciones como la selección de voz y ajustes visuales. También se evalúa la posibilidad de integrar modelos de aprendizaje automático para analizar el comportamiento del usuario y anticipar sus necesidades comunicativas, con el objetivo de ofrecer una experiencia aún más adaptada y proactiva.

Estas mejoras garantizarán que el proyecto continúe evolucionando y adaptándose a las necesidades cambiantes de los usuarios con TEA.

Referencias

1. CDC. (s.f.). Prevalencia del autismo más alta, según los datos de 11 comunidades de la Red de ADDM. Obtenido de Prevalencia del autismo más alta, según los datos de 11 comunidades de la Red de ADDM: https://www.cdc.gov/spanish/mediosdecomunicacion/comunicados/p_autismo_032323.html.
2. D. Graffox, «IEEE Citation Reference,» 2009. [En línea]. Available: <http://www.ieee.org/documents/ieeecitationref.pdf>, «"Los trastornos del espectro autista (TEA) son discapacidades del desarrollo",» [En línea]. Available: <https://www.cdc.gov/ncbddd/spanish/autism/signs.html>.
3. CDC, «Trastornos del espectro autista (TEA),» 15 Mayo 2024. [En línea]. Available: <https://www.cdc.gov/autism/es/about/sobre-los-trastornos-del-espectro-autista.html>.
4. A. i. (i. advance), «Qué es el autismo, y qué lo causa,» 10 Diciembre 2021. [En línea]. Available: <https://www.adninstitut.com/que-es-el-autismo-y-que-lo-causa-n-52-es>.
5. F. ConectaTEA, «Sistemas alternativos y aumentativos de comunicación (SAAC),» 17 Junio 2020. [En línea]. Available: <https://www.fundacionconectea.org/2020/06/17/sistemas-alternativos-y-aumentativos-de-comunicacion-s-a-a-c/>.
6. CDC, «Prevalencia del autismo más alta, según los datos de 11 comunidades de la Red de ADDM,» [En línea]. Available: https://www.cdc.gov/spanish/mediosdecomunicacion/comunicados/p_autismo_032323.html.

7. N. I. o. D. a. O. C. Disorders, «Trastornos del espectro autista: Problemas de comunicación en los niños,» 24 abril 2020. [En línea]. Available: <https://www.nidcd.nih.gov/es/espanol/problemas-de-comunicacion-en-los-ninos-con-trastornos-del-espectro-autista#2>.
8. C. H. d. Toro, «TRABAJO FINAL DE GRADO EN MAESTRO/A DE EDUCACIÓN PRIMARIA LA EDUCACIÓN INCLUSIVA, ¿UNA ESCUELA PARA TODOS?,» 2016.
9. RACO, «Aportaciones de las tecnologías de apoyo en personas con autismo,» [En línea]. Available: <https://www.raco.cat/index.php/AmbitsAAF/article/download/370024/463738/>.
10. Isabel del Hoyo (AETAPI), «LA ÉTICA EN EL APOYO A LAS PERSONAS CON TEA: IMPLICACIONES PRÁCTICAS,» 16 Mayo 2018. [En línea]. Available: <https://aetapi.org/la-etica-en-el-apoyo-a-las-personas-con-tea/>.
11. ACM, «Association for Computing Machinery,» [En línea]. Available: <https://www.acm.org/>.
12. Wikipedia, «Association for Computing Machinery,» [En línea]. Available: https://es.wikipedia.org/wiki/Association_for_Computing_Machinery.
13. ACM, «ACM Code of Ethics and Professional Conduct,» [En línea]. Available: <https://www.acm.org/code-of-ethics>.
14. [14]C. E. Navarra, «AsTeRICS GRID,» 24 noviembre 2022. [En línea]. Available: <https://creena.educacion.navarra.es/web/recursos/2022/11/24/asterics-grid/>.
15. C. C. A. d. R. p. I. E. y. I. I. (. d. Aragon), «ARASAAC,» [En línea]. Available: <https://carei.es/descipcion-ayasaac/#:~:text=ARASAAC%20es%20un%20Sistema%20Aumentativo,%2C%20traumatismos%20y%20degeneraci%C3%B3n%20cognitiva>.
16. M. J. G. M. C. E. Baldassarri S., «AraBoard: A Multiplatform Alternative and Augmentative Communication Tool In proceedings of 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion,» DSAI, Vigo, Spain, 2013.
17. M. J. C. E. M. L. Baldassarri S., «Accessibility Evaluation of an Alternative and Augmentative Communication (AAC),» LNCS 8516, 2014.
18. S. H. R. H.-G. A. M.-D. C. P. S. G. P. Bautista, «AraTraductor: Text to Pictogram Translation using Natural Language Processing Techniques,» 19 Junio 2017. [En línea]. Available: <http://nil.fdi.ucm.es/?q=node/680>.
19. H. R. B. S. a. G. P. G. Mendez, «PictoEditor: un editor predictivo basado en pictogramas para facilitar la interacción con personas con discapacidad,» 2018. [En línea]. Available: <http://nil.fdi.ucm.es/sites/default/files/Pictoeditor-Interaccion2018.pdf>.
20. D. M.-E. I. M. J. M. L. H. R. B. S. Guasch, «Expert-based Assessment of an Augmentative and Alternative Communication Tool,» ACM Digital Library, Donostia, Gipuzkoa, Spain, 2019.

KetoGlut Planner: Aplicación de Ayuda para Personas con Deficiencia de Glut-1

Ana Covadonga Schleissner¹, Susana Bautista¹

¹ Escuela Politécnica Superior, Universidad Francisco de Vitoria, Madrid, España.
7904007@alumnos.ufv.es, susana.bautista@ufv.es

Abstract. En este proyecto se diseña una aplicación web con el objetivo de mejorar la calidad de vida de las personas que padecen una enfermedad rara denominada Síndrome de Déficit de Glut-1. Se ha llevado a cabo siguiendo el "Diseño Centrado en el Usuario", involucrando en su desarrollo a madres voluntarias de la asociación asGLUTdiece. Esta asociación está formada por familias cuyo objetivo es difundir información sobre la enfermedad entre hospitales, centros de investigación, terapeutas, pacientes y familiares. Los requisitos de la aplicación web fueron recogidos por expertos en la enfermedad que ayudaron a definir sus funcionalidades. Se creó un primer prototipo y se probó por ellos para rediseñar la aplicación, basándose en sus comentarios y mejoras. El trabajo presenta el diseño, la implementación y la evaluación realizada a lo largo del proyecto. Los resultados de esta evaluación muestran una gran aceptación de la aplicación desarrollada por parte de los usuarios finales, ayudándoles a interactuar con las dietas necesarias para las personas diagnosticadas con Síndrome de Déficit de Glut-1. Este trabajo contribuye al área de conocimiento al presentar un estudio de caso exitoso de diseño participativo aplicado a herramientas de apoyo a enfermedades raras, demostrando cómo el desarrollo colaborativo con las partes interesadas puede conducir a soluciones de salud digital impactantes.

Keywords: Enfermedad Rara, Glut-1, Aplicación Web, Cetosis, Diseño Centrado en el Usuario.

1 Introducción

El síndrome de deficiencia del transportador de glucosa cerebral de tipo 1 (Glut-1) es una rara condición genética que altera el metabolismo cerebral. Se han notificado 500 casos en el mundo. Está causada por mutaciones o variantes en el gen SLC2A1, que codifica la proteína transportadora de glucosa tipo 13. Como consecuencia de esta mutación, la enfermedad se caracteriza por un impedimento del transporte de glucosa al cerebro, ya que Glut-1 es su principal transportador, aunque también mueve otros

azúcares importantes a través de la barrera hematoencefálica. En condiciones normales, la glucosa es la fuente principal de energía del cerebro, es por ello por lo que las mutaciones en este gen pueden reducir o eliminar la función del transportador y provocar síntomas graves.

La alteración del transporte de glucosa asociada con el Síndrome de Deficiencia de Glut-1 crea una crisis metabólica en el cerebro y a menudo resulta en una variedad de síntomas neurológicos. Su forma clásica (85% de casos) se manifiesta por crisis epilépticas de inicio temprano y refractarias a tratamiento, microcefalia, retraso del neurodesarrollo, déficit intelectual, trastorno del lenguaje y trastornos en el movimiento. No todos los pacientes experimentan todos los síntomas y existe un amplio espectro de gravedad. Los síntomas pueden cambiar y evolucionar con el tiempo.

El tratamiento estándar de atención es una dieta cetogénica. Esta dieta baja en carbohidratos, moderada en proteínas y alta en grasas hace que el cuerpo produzca cetonas, las cuales el cerebro y otros tejidos pueden utilizar como combustible cuando la glucosa es limitada. La dieta ayuda a la mayoría de los pacientes con la mayoría de los síntomas, incluso en la edad adulta, pero no es una cura. Sin embargo, se cree que las dietas cetogénicas ayudan a preservar el crecimiento y desarrollo del cerebro en personas con el Síndrome de Deficiencia de Glut-1, por lo que el diagnóstico y tratamiento tempranos son fundamentales.

Actualmente no existe cura para el Síndrome de Deficiencia de Glut-1. El tratamiento tiene como objetivo controlar los síntomas abordando lo que se cree que es el problema fundamental del Síndrome de Deficiencia de Glut-1: una crisis metabólica en el cerebro debido a un transporte deficiente de glucosa a través de la barrera hematoencefálica. El estándar de atención actual es una dieta cetogénica supervisada por un médico.

Las dietas cetogénicas médicas se adaptan cuidadosamente a cada paciente, tienen posibles efectos secundarios y sólo deben usarse bajo la supervisión de profesionales médicos. Las dietas cetogénicas pueden ayudar a mejorar la mayoría de los síntomas asociados con el Síndrome de Deficiencia de Glut-1, pero a menudo no los controlan por completo. Actualmente, los investigadores del Síndrome de Deficiencia de Glut-1 están investigando otras formas de tratamiento, pero hasta el momento no hay otros tratamientos disponibles aprobados específicamente para tratar los mecanismos subyacentes del Síndrome de Deficiencia de Glut-1.

Puede pasar algún tiempo hasta que todo el potencial de la dieta surta efecto y, en ocasiones, es necesario realizar ajustes y modificaciones en los parámetros de la dieta (proporción, calorías, niveles de proteínas, número y horario de las comidas, fuentes de grasas, etc.).

El diagnóstico temprano es de vital importancia para que la dieta pueda iniciarse lo antes posible y así poder ayudar a controlar los síntomas y preservar mejor el crecimiento y desarrollo del cerebro, además, esto evitaría ensayos de medicamentos y efectos secundarios a menudo ineficaces y potencialmente dañinos. Los expertos recomiendan continuar la dieta el mayor tiempo posible y muchos pacientes continúan experimentando los beneficios de la dieta incluso hasta la edad adulta.

La propuesta para asistir a las personas que padecen esta enfermedad consiste en el desarrollo de una aplicación que facilite la creación de dietas personalizadas. Además, la aplicación incluirá un listado de alimentos con su información nutricional, así como una selección de recetas que contribuirán a seguir adecuadamente la dieta requerida.

2 Diseño de la aplicación KetoGlut Planner

Para el desarrollo de la aplicación, se ha adoptado una arquitectura Modelo-Vista-Tem-plate (MVT), en la cual el cliente es el navegador web del usuario, mientras que el servidor Django se encarga de procesar las solicitudes y proporcionar las respuestas necesarias. Para el alojamiento en la nube, se ha optado por un servidor privado.

El servidor privado VPS, adquirido en OVH Cloud16, utiliza el sistema operativo Ubuntu y el servidor web Apache2. Para el despliegue de la aplicación, el servidor VPS ha sido configurado con Ubuntu, instalando y ajustando Apache2 para gestionar las solicitudes HTTP. Además, se ha configurado un subdominio específico en Apache2, modificando los archivos de configuración correspondientes para garantizar que el contenido de la web se sirva correctamente desde dicho subdominio.

En el lado del frontend, se han empleado Django, TailwindCSS y JavaScript para el desarrollo de las páginas web. Django estructura el contenido de las páginas, TailwindCSS proporciona estilo y diseño a los elementos visuales, y JavaScript se utiliza para añadir interactividad y funcionalidad a la aplicación. Todo ello complementado con las herramientas que Django ofrece para la integración del frontend con el backend.

Para el backend, se ha utilizado Python como lenguaje de programación principal y Django como framework web. Django, por su robustez y escalabilidad, permite el desarrollo ágil de aplicaciones web y API (Application Programming Interface) gracias a sus numerosas funcionalidades integradas. Se han creado rutas y vistas en Django para gestionar las solicitudes del cliente, procesar la lógica del servicio y acceder a la base de datos mediante su ORM (Object-Relational Mapping).

Además, se han aprovechado características como el sistema de autenticación, la administración automática de interfaces y el manejo de sesiones para garantizar un funcionamiento eficiente y seguro.

En el desarrollo y despliegue de la aplicación, se han utilizado tecnologías modernas y potentes. En el frontend, Django, TailwindCSS y JavaScript han permitido la creación de páginas web atractivas y funcionales. En el backend, Django ha sido el framework principal, proporcionando herramientas avanzadas para la gestión

de rutas, vistas y acceso a la base de datos. Este enfoque ha permitido construir una aplicación robusta y eficiente, adaptada a las necesidades del proyecto y optimizada para ofrecer una experiencia de usuario óptima.

2.1 Implementación de KetoGlut Planner

Para la implementación de la aplicación web, se ha optado por un enfoque de gestión de proyectos eficaz y bien organizado, abarcando detalladamente las siguientes etapas del proceso de implementación como se pueden ver en la Figura 1 y se detalla a continuación:

1. **Definición de objetivos:** Se definen los requisitos requeridos por los usuarios en términos de diseño, funcionalidad y usabilidad. Se tuvo en cuenta que tiene que ser una aplicación web accesible para todos los usuarios.
2. **Planificación:** En esta fase, se planifica como va a ser el desarrollo, tanto los Frameworks que se van a utilizar como los tiempos y recursos de las que se van a disponer en las siguientes fases del proyecto.
3. **Desarrollo del backend y base de datos:** Para el desarrollo del backend se ha empleado el framework de Django con el lenguaje de programación Python conectado con la base de datos en MySQL Lite.
4. **Desarrollo del frontend:** El desarrollo del frontend se ha realizado utilizando las tecnologías Django, TailwindCSS y JavaScript para hacer el diseño de la página web para que ésta sea accesible para todos los usuarios.
5. **Implementación y despliegue:** En esta fase se despliega la aplicación en un servidor privado VPS adquirido en OVH Cloud para así poder ser accesible a los usuarios finales de la aplicación web.
6. **Depuración de errores:** Se corrigen errores básicos encontrados y se realizan diferentes cambios necesarios para mejorar la aplicación web.

La implementación se ha realizado con éxito, cubriendo todos los requisitos que los usuarios de la asociación nos pidieron para la creación de recetas y menús que cumplen con los requisitos de la dieta cetogénica.

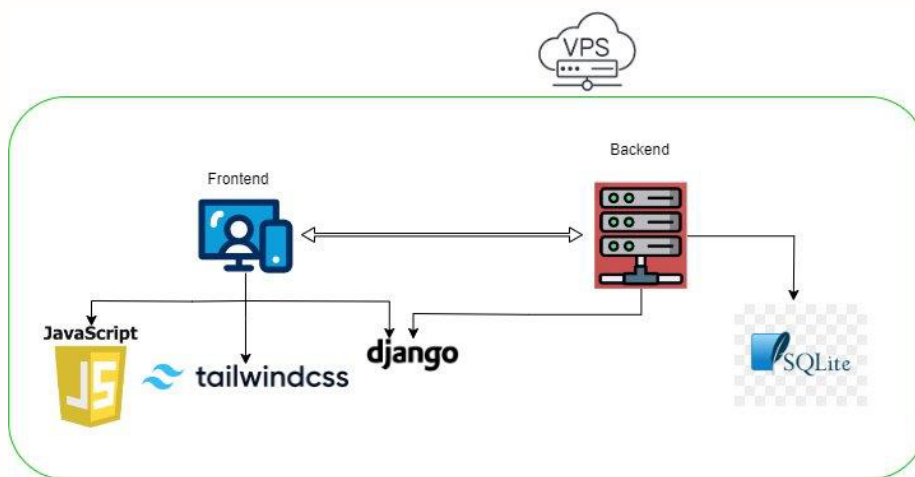


Fig. 1. Esquema de la implementación. KetoGlut Planner 5.

3 Evaluación

La primera evaluación que se llevó a cabo fue la muestra del primer prototipo que se implementó en AdobeXD. Esta presentación y evaluación se hizo en una reunión con una de las madres de la asociación. En esta reunión se comentaron distintas funcionalidades extra que debía tener la aplicación web para que estuviera más completa. Tras esta reunión, se implementaron los cambios y se realizó el diseño final. Una vez terminado el diseño final se empezó con la implementación de la aplicación web.

Cuando se finalizó el desarrollo de la aplicación web, los usuarios finales hicieron uso de la misma. Se incluyó un manual de usuario para que pudieran hacer uso de todas las funcionalidades que ofrece KetoGlut Planner, y así poder contestar a todas las preguntas de los formularios. Tras una semana explorando la aplicación y probando todas sus funcionalidades, se enviaron los formularios para poder analizar el cumplimiento de los objetivos y su usabilidad.

Se ha optado por hacer uso del formulario SUS para evaluar su usabilidad desde la perspectiva de los usuarios finales. Este tipo de pacientes, así como sus familiares o cuidadores, requieren una plataforma intuitiva, eficiente y accesible, ya que de ello depende en parte la correcta implementación de una dieta terapéutica fundamental para su salud. El SUS permite obtener una medida estandarizada y rápida de la experiencia del usuario, facilitando la identificación de áreas de mejora en la interfaz y funcionalidad de la aplicación. Además, al tratarse de un cuestionario validado internacionalmente, sus resultados ofrecen una base objetiva para comparar versiones futuras de la herramienta, asegurando que la solución final esté verdaderamente centrada en las necesidades de quienes la utilizarán diariamente

3.1 Formulario SUS

Tabla 1. Formulario SUS.

Después de terminar las tareas, por favor conteste las siguientes preguntas de usabilidad mediante este Cuestionario de usabilidad:	
1.- Totalmente en desacuerdo	
2.- En desacuerdo	
3.- Neutro	
4.- De acuerdo	
5.- Totalmente de acuerdo	
Me gustaría usar esta herramienta frecuente-mente	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero que esta herramienta es innecesaria-mente compleja	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero que la herramienta es fácil de usar	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero necesario el apoyo de personal experto para poder utilizar esta herramienta	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero que las funciones de la herramienta están bien integradas	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero que la herramienta presenta muchas contradicciones	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Imagino que la mayoría de las personas aprenderían a usar esta herramienta rápidamente	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Considero que el uso de la herramienta es tedioso	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Me sentí muy confiado al usar la herramienta	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo
Necesité saber bastantes cosas antes de poder empezar a usar esta herramienta.	1- Totalmente en desacuerdo
	5- Totalmente de acuerdo

3.2 Formulario sobre la aplicación web

Tabla 3. Formulario sobre la aplicación web.

Una vez que has usado la aplicación KetoGlut Planner, por favor, responde a las siguientes preguntas. Donde	
1- es Totalmente en desacuerdo	
5- es Totalmente de acuerdo	
Gracias por participar y acompañarme en el Proyecto Final de Grado.	

Creo que el uso de KetoGlut Planner me permitirá crear un menú más fácilmente.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que el uso de KetoGlut Planner reduciría el tiempo necesario para generar un menú.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la página de inicio de KetoGlut Planner me da una idea global de lo que es la enfermedad de déficit de Glut-1.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la disposición de los alimentos en KetoGlut Planner me ayuda a visualizarlos de manera más efectiva.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que el cambio en la disposición de los alimentos en KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la disposición de las recetas en KetoGlut Planner me ayuda a visualizarlos de manera más efectiva.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que el cambio en la disposición de las recetas en KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que el cambio en la disposición de las recetas en KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de poder filtrar las recetas por ratio de KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de saber las recetas que contiene cada alimento de KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de poder agregar y modificar recetas de KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de poder agregar y modificar alimentos de KetoGlut Planner es útil.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de poder ver las recetas creadas por todos los usuarios de KetoGlut Planner me dará más posibilidades en la dieta.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que la funcionalidad de poder ver los alimentos agregados por todos los usuarios de KetoGlut Planner me dará más posibilidades en la dieta.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo
Creo que el uso de KetoGlut Planner hará que me sienta más conectada con la gente en mi misma situación.	1- Totalmente en desacuerdo 5- Totalmente de acuerdo

3.3 Análisis de resultados de los formularios

Tras haber analizado los resultados obtenidos de 11 usuarios finales que han podido hacer uso de la aplicación web, KetoGlut Planner recibe en general una valoración muy positiva por su parte. Es altamente valorado por su facilidad para crear menús y por reducir el tiempo empleado en ello, con puntuaciones casi perfectas en estos aspectos. La herramienta también es bien valorada por su capacidad para proporcionar información educativa sobre la enfermedad de déficit de Glut-1 y por ser rápida a la hora de aprender a utilizarla.

No obstante, hay una percepción de tediosidad y una necesidad de conocimientos previos que varía entre los usuarios, lo cual indica que en un futuro habría que tener una reunión con los usuarios finales para analizar cómo poder mejorar estos aspectos. Una posible solución planteada es considerar simplificar aún más la interfaz de usuario y proporcionar más información sobre la aplicación, para que los usuarios que visiten la aplicación web y que no estén familiarizados con la enfermedad, puedan entenderla mejor.

Para poder obtener el resultado del cuestionario SUS, habrá que sumar las respuestas de los enunciados impares y después restar 5, después se sumará la respuesta de los enunciados pares y restar ese total a 25. Una vez tengamos los resultados, lo sumaremos y multiplicaremos por 2,5. Tras hacer los cálculos de los resultados del cuestionario SUS, siguiendo la fórmula correspondiente, KetoGlut Planner ha recibido una puntuación promedio de 71,82. Al estar entre los valores 68 y 100 podemos ver que la usabilidad de la aplicación web es aceptable.

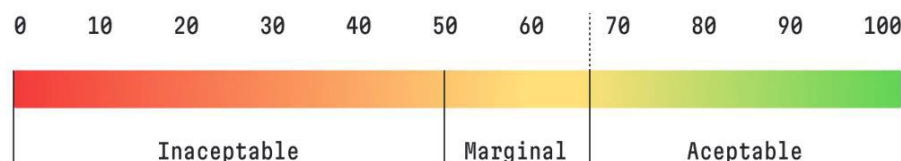


Fig. 2. Escala de resultados encuesta SUS.

En conclusión, KetoGlut Planner es una herramienta eficaz y bien recibida en la mayoría de los aspectos, con un fuerte potencial para ser mejorada y optimizada en áreas clave para proporcionar una experiencia de usuario aún más satisfactoria y accesible. Se han recogido varias sugerencias de los usuarios finales que han podido hacer uso de ésta para futuras mejoras ya que hay mucho interés en seguir trabajando en este proyecto porque es una enfermedad poco conocida y que necesita de la tecnología para facilitar la vida a sus usuarios y a sus familias.

4 Conclusiones y trabajo futuro

El desarrollo de la aplicación web KetoGlut Planner, diseñada para generar dietas específicas para personas con Síndrome de Deficiencia de glut-1, ha permitido no solo la aplicación práctica de los conocimientos adquiridos a lo largo del grado en Ingeniería Informática, sino también una inmersión profunda en un ámbito sanitario complejo y poco explorado, que presenta desafíos únicos tanto a nivel técnico como humano.

Uno de los principales retos al inicio del proyecto fue el desconocimiento total sobre la enfermedad, lo que implicó un proceso intensivo de documentación, formación autodidacta y aprendizaje constante sobre el Síndrome de Deficiencia de glut-1 y sus implicaciones nutricionales. A ello se sumó la dificultad de establecer contacto con los usuarios finales, lo cual fue fundamental para entender sus necesidades reales y garantizar que la solución propuesta tuviera un impacto tangible en su calidad de vida. Gracias a la colaboración estrecha con la asociación asGLUTdiece y al compromiso continuo con sus recomendaciones y sugerencias, fue posible evolucionar el proyecto hasta desarrollar una herramienta adaptada, accesible y útil.

Los resultados obtenidos han sido altamente favorables. Se ha cumplido no solo con los objetivos técnicos establecidos al inicio del proyecto, sino también con el objetivo final y más importante: ayudar a personas que conviven con el síndrome de deficiencia de glut-1, ofreciéndoles una solución que facilita la planificación de sus dietas y la creación de recetas compatibles con su tratamiento. Las pruebas realizadas con usuarios reales, así como las respuestas positivas obtenidas a través del cuestionario de usabilidad SUS, confirman que la herramienta no solo es funcional, sino también eficaz y valorada por quienes la necesitan.

Este proyecto reafirma la importancia de la interdisciplinariedad, ya que la combinación de conocimientos en ingeniería informática con comprensión médica y sensibilidad social ha sido clave para su éxito. Asimismo, resalta el valor de diseñar con empatía, comprendiendo a fondo las circunstancias del usuario final, lo que ha guiado decisiones cruciales de diseño e implementación.

A nivel personal y profesional, esta experiencia ha abierto nuevas perspectivas en el campo de la tecnología aplicada a la salud, despertando un interés particular por la innovación en el ámbito sanitario. La satisfacción de ver cómo una solución tecnológica puede mejorar la vida diaria de personas reales ha sido profundamente motivadora. Además, enfrentar desafíos como la falta de información inicial y la necesidad de adquirir competencias técnicas adicionales ha fortalecido no solo las habilidades de desarrollo, sino también la resiliencia, la capacidad de adaptación y la vocación de servicio, consolidando una base sólida para futuros proyectos con impacto social.

A lo largo del desarrollo del proyecto KetoGlut Planner, se han identificado diversas áreas de mejora y expansión que podrían incrementar significativamente el valor y la utilidad de la aplicación. A continuación, se detallan los trabajos futuros que se proponen para continuar con el desarrollo y perfeccionamiento de la aplicación:

1. Inclusión de Médicos y Creación de una Vista Exclusiva para Profesionales de la Salud

Una de las prioridades para el futuro desarrollo de KetoGlut Planner es la inclusión de médicos en el proceso de gestión y planificación dietética. La creación de una vista específica para profesionales de la salud permitirá a los médicos tener acceso a funcionalidades avanzadas que no están disponibles para los usuarios finales. Esta vista médica incluirá herramientas para monitorizar la evolución de los pacientes, ajustar las dietas basadas en resultados clínicos y gestionar recetas de manera más efectiva. Al proporcionar una interfaz dedicada, se facilitará una mejor colaboración entre médicos y pacientes, asegurando que las dietas recomendadas sean más precisas y personalizadas según las necesidades individuales de cada paciente.

2. Entrenamiento de una Inteligencia Artificial para la Generación Automática de Dietas

Otra área de desarrollo crucial es el entrenamiento de una inteligencia artificial (IA) que permita la generación automática de dietas. Actualmente, la aplicación utiliza una ketocalculadora que requiere la introducción manual de datos y parámetros por parte del usuario. La implementación de una IA podría automatizar este proceso, generando dietas personalizadas de manera eficiente y precisa sin necesidad de intervención manual. Esta IA se entrenaría utilizando datos históricos y patrones dietéticos efectivos, mejorando continuamente a través de la retroalimentación proporcionada por los usuarios y médicos. Esta innovación no solo ahorraría tiempo, sino que también podría aumentar la precisión y adecuación de las dietas recomendadas.

3. Ampliación de la Base de Datos con Más Recetas y Alimentos

Para mejorar la utilidad y atractivo de la aplicación, es esencial ampliar la base de datos actual de recetas y alimentos disponibles. Incluir una mayor variedad de recetas y una lista más extensa de alimentos permitirá a los usuarios disponer de opciones más diversas y adaptadas a sus gustos y necesidades dietéticas. Este proceso de ampliación requerirá la colaboración con nutricionistas y expertos en dietética para asegurar que las nuevas recetas y alimentos cumplen con los requisitos necesarios para el manejo del Síndrome de Deficiencia de Glut-1. Además, se podrían incorporar funcionalidades que permitan a los usuarios sugerir y compartir sus propias recetas, creando una comunidad colaborativa que enriquezca el contenido de la aplicación y la sensación de pertenencia.

Estos trabajos futuros representan un conjunto de mejoras y expansiones que no solo incrementarán la funcionalidad de KetoGlut Planner, sino que también mejorarán significativamente la experiencia del usuario y la efectividad de la herramienta. La inclusión de médicos y la creación de una vista exclusiva para ellos, el entrenamiento de una IA para la generación automática de dietas y la ampliación de la base de datos con más recetas y alimentos son pasos fundamentales para asegurar que la aplicación continúe evolucionando y adaptándose a las necesidades cambiantes de sus usuarios. Con estas mejoras, KetoGlut Planner estará mejor equipada para cumplir su misión de mejorar la calidad de vida de las personas con Síndrome de Deficiencia de Glut-1, proporcionando una herramienta integral y avanzada para la gestión de sus dietas.

Referencias

1. D. Graffox, «IEEE Citation Reference,» 2009. [En línea]. Available: <http://www.ieee.org/documents/ieeecitationref.pdf>.
2. «Federación Española de Enfermedades Raras,» 2022. [En línea]. Available: <https://www.enfermedades-raras.org/enfermedades-raras>.
3. M. Zabala, V. Castillo, M. González y S. Castillo. "Síndrome de deficiencia de GLUT-1". Neurologia.com. Accedido el 16 de febrero de 2024. [En línea]. Disponible: <https://neurologia.com/articulo/2019126>.
4. Colaboradores de los proyectos Wikimedia. "GLUT1 - Wikipedia, la enciclopedia libre". Wikipedia, la enciclopedia libre. Accedido el 8 de junio de 2024. [En línea]. Disponible: <https://es.wikipedia.org/wiki/GLUT1>.

5. "¿Qué es el GLUT1 y dónde se encuentra?" Guía metabólica. Accedido el 8 de junio de 2024. [En línea]. Disponible: <https://metabolicas.sjdhospitalbarcelona.org/ecm/deficiencia-transportador-glucosa-glut-1/info/es-glut1-donde-encuentra>.
6. C. Gómez Lado, M.L. Couce Pico, A. Sánchez-Salmón y J.M. Pascual. "Deficiencia del transportador de glucosa tipo I: una enfermedad neuro metabólica tratable | Anales de Pe-diatría". Anales de Pediatría. Accedido el 8 de febrero de 2024. [En línea]. Disponible: <https://www.analesdepediatria.org/es-deficiencia-del-transportador-glucosa-tipo-articulo-S1695403308720654>.
7. F. Arrieta Blanco, A. Bélanger Quintana, N. Bengoa Rojano, S. Stanesco y M. Martínez Pardo. "Síndrome de deficiencia de GLUT1 y COVID-19". PubMed Central (PMC). Accedido el 8 de febrero de 2024. [En línea]. Disponible: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7462458/>
8. Colaboradores de los proyectos Wikimedia. "Cetosis - Wikipedia, la enciclopedia libre". Wikipedia, la enciclopedia libre. Accedido el 9 de febrero de 2024. [En línea]. Disponible: <https://es.wikipedia.org/wiki/Cetosis>.
9. P. Canal. "¿Qué es el diseño centrado en el usuario?" Thinking for Innovation. Accedido el 2 de marzo de 2024. [En línea]. Disponible: [https://www.iebschool.com/blog/disenio-cen-trado-en-el-usuario-analitica-usabilidad/#:~:text=El%20diseño%20cen-trado%20en%20el%20usua-rio%20\(DCU\)%20es%20una%20metodología,la%20conceptualización%20hasta%20la%20implementación](https://www.iebschool.com/blog/disenio-cen-trado-en-el-usuario-analitica-usabilidad/#:~:text=El%20diseño%20cen-trado%20en%20el%20usua-rio%20(DCU)%20es%20una%20metodología,la%20conceptualización%20hasta%20la%20implementación).
10. Iberdrola. "DISEÑO CENTRADO EN EL USUARIO". Iberdrola. Accedido el 2 de marzo de 2024. [En línea]. Disponible: <https://www.iberdrola.com/innovacion/que-es-diseno-cen-trado-en-el-usuario>.
11. P. Bermudez Talavera. "Diseño Centrado en el Usuario y las etapas de su metodología". Pedro Bermudez Talavera. Accedido el 2 de marzo de 2024. [En línea]. Disponible: <https://pedrobermudeztalavera.com/disenio-centrado-en-el-usuario-y-las-etapas-de-su-metodologia/>.
12. S. Pursell. "Diseño centrado en el usuario: qué es, etapas y ejemplos". Blog de HubSpot | Marketing, Ventas, Servicio al Cliente y Sitio Web. Accedido el 2 de marzo de 2024. [En línea]. Disponible: <https://blog.hubspot.es/website/disenio-centrado-usuario>.
13. Colaboradores de los proyectos Wikimedia. "Diseño centrado en el usuario - Wikipedia, la enciclopedia libre". Wikipedia, la enciclopedia libre. Accedido el 2 de marzo de 2024. [En línea]. Disponible: https://es.wikipedia.org/wiki/Diseño_centrado_en_el_usuario#:~:text=El%20Diseño%20Centrado%20en%20el,mínimo%20es-fuerzo%20de%20su%20parte.

Multiplatform System for the Prediagnosis of Hearing Loss: Implementation and Validation

Diany Isabel García-Carrascal¹[0009-0001-9320-2124], Jose Leonardo Quintero-Leon¹[0009-0002-6350-5674],
Magreth Rossio Sanguino-Reyes¹[0000-0002-2194-7485], Huizilopoztli Luna-García²[0000-0001-5714-7482],
Dewar Rico-Bautista¹[0000-0002-1808-3874]

¹ Universidad Francisco de Paula Santander, Ocaña, Colombia

² Laboratorio de Tecnologías Interactivas y Experiencia de Usuario (LITUX), Universidad Autónoma de Zacatecas,
Zacatecas, México.

{digarciac, jlquinterol, mrsanguinor, dwricob}@ufpso.edu.co, hlugar@uaz.edu.mx

Abstract. Hearing loss in Colombia, according to the Ministry of Health, translates into five million people suffering from hearing loss, which represents 11% of the population. However, only three out of ten people seek help and go to a specialist, translating into a low early detection of hearing diseases. The purpose of this paper is to present the design, development, and validation of multiplatform accessibility and usability. In this sense, a review of the techniques and tools used to diagnose hearing loss is performed. It is concluded that the use of digital platforms and applications can help to perform an early prediagnosis of hearing diseases. The context of the proposed solution is presented: a prototype consisting of a multiplatform system that seeks to obtain a pre-diagnosis of the disease.

Keywords: Hearing Impairment, Hearing Loss, Prediagnosis, Prediagnostic Techniques.

1 Introduction

Hypoacusis and how it is generated, according to [1, 2] hypoacusis is defined as "the partial or total loss of the ability to perceive or understand sound", among the types of hypoacusis that a person may face are, according to [3] congenital hypoacusis, also sudden hypoacusis [3], a different kind of hypoacusis according to [3] is conductive. On the other hand [3], the disease is caused due to the large number of genes involved, with a total of 120, also shows a classification of the results obtained in which the family medical history and their inheritance patterns are evidenced, referring that hearing loss is hereditary, but this is not the only case, taking into account [4, 5]: "Hearing impairment, or hearing loss, can be caused by several factors and the implications vary according to the type, degree, cause and age of onset".

According to this and observing the difficulties that are presented to this type of people from an early age, as shown in a study by [6] "Hearing impaired children often have delayed language comprehension" this in turn presents a possible solution, which was: "to present and develop a new application based on assistive technology on the Android platform to help people with speech disabilities and HI".

Our case only focuses on an adult population considering that these people are the ones who will make use of our research project in which we seek to facilitate the pre-diagnosis of hearing loss and appointment booking management for people with this disability, thus designing a multiplatform web-oriented system, although according to [5], "there are still some obstacles by developers, such as lack of interest, motivation and knowledge about how to properly develop an accessible web application" for people with this problem, motivation and necessary knowledge on how to properly develop an accessible web application" for people with this problem, being us and through research, the cause of breaking the gap that has been generated according to the passage of time, in turn also give some results of systematic mappings of how they are designing and implementing these web interfaces, of which not many tools that help developers were found.

Therefore, we decided to investigate more about it until we found a paper that designed an Android application with a similar objective to ours: "to present an application designed specifically for deaf or hard of hearing people, to help them understand what other people tell them". [7] In our case, as mentioned above, a

web application will be designed to manage a hearing impairment pre-diagnosis system and appointment booking.

According to [8] their article, "Appointment Booking in Customer Choice," two models are shown: the non-sequential offer and the sequential offer. By comparing these models, the study generated especially useful information for improving current appointment booking processes, with the sequential model being considered the most optimal. "Our extensive numerical study shows that sequential bidding can significantly improve the slot occupancy rate (by 6% to 8% on average, and up to 18% in our test cases) compared to non-sequential bidding. The time slots refer to the spaces available for the customer to book."

With the great boom in technology, most organizations (if not all) use tools that help optimize their processes, something that clinics and hospitals must handle to a greater extent since these entities are responsible for providing medical care to people around the world [9, 10]. Because of this, many innovative ideas are constantly being generated that help to improve the functioning of the processes of organizations that offer health services, all to optimize the activities performed by specialists on patients, bringing benefits for both parties as time and resources are optimized [11].

Considering that hearing loss is one of the most widespread disabilities and should be addressed by early detection and intervention, there is currently a great interest in cost-effective and massively deployable hearing screening technologies, where smartphones and other electronic devices are ideal for hearing loss screening [12]. According to [13], to perform an audiometer hearing test patients must go to a hospital to be examined and evaluated by an audiologist in a soundproof room, where in some cases this type of audiometer testing may not be convenient or able to be performed because not all hospitals have an audiometer or an audiologist available. Therefore, performing technological tests through a digital platform using a telephone or electronic device is of vital importance for this type of situation, providing a solution so that these hearing loss screening tests can be performed and evaluated by specialists in the area.

Given the difficulty of performing early prediagnosis tests for patients with hearing loss and the need to go to a hospital to perform them, it is proposed to develop a multiplatform system that allows the prediagnosis of people with hearing loss through an electronic device and from anywhere. It should also include a database of specialists with whom the patient can communicate to schedule an appointment, according to the pre-diagnosis of the platform and the need to go to a medical professional.

According to [14] context is needed for "Data were collected on personalities diagnosed with hearing loss in the 4th century B.C.". However, it was not until the 1940s that the term "hearing loss" began to be used in publications, as well as the term "audiologist." Audiology is the medical specialty concerned with prevention, diagnosis, and treatment, and its origins are said to date back to the 19th century. It has also been said that the first known deaf man was at the time of the Roman Empire when having hearing loss was not synonymous with an easy life.

The document is organized as follows: first, the introduction, where the problematic situation is presented; second, the methodology, where the design of the research process to obtain and evaluate the proposed solution is presented; and finally, the results, together with their respective discussion.

2 Methodology

The research approach used was quantitative since it refers to the approach to reality understood as objective and considers that the researcher must distance him/herself from it to analyze it [15]. From this approach, the researcher considers that data should be discovered and analyzed objectively [16] so that phenomena in a population can be described, measured, explained, and predicted. The scope was descriptive, as it allowed the development of a quantitative approach: "The objective of descriptive research is to learn about situations, customs and prevailing attitudes through the accurate description of activities, objects, processes, and people". On the other hand, it allowed the collection of quantifiable information for use in the statistical analysis of the population sample, from the quantitative approach [17, 18].

2.1 Population and sample

For the implementation of this project, the population will be composed of students of the Systems Engineering program of the Universidad Francisco de Paula Santander Ocaña, with an average age between 18 and 27 years old, who are currently studying. The following formula was used to obtain the sample data:

$$n = \frac{z^2(p \times q)}{e^2 + \frac{(z^2(p \times q))}{N}} \quad (1)$$

Where n is the sample size, z is the desired confidence level, p is the proportion of the population with the desired characteristic, q is the proportion of the population without the desired characteristic, the level of error willing to commit, and N is the population size. For the values of p and q, given the lack of previous data, a conservative assumption was made, where p = q = 0.5, assuming that half of the population has the desired characteristic and half does not. For the z value, a value of 1.96 was taken, since it corresponds to the 97.5 percentile of the normal distribution, which is used to calculate confidence intervals for a 95 % level. The following formula was used:

$$\begin{aligned} n &= \frac{1.96^2(0.5 \times 0.5)}{0.10^2 + \frac{(1.96^2(0.5 \times 0.5))}{350}} \\ n &= \frac{(3.8416 * 0.25)}{0.01 + \frac{(3.8416 * 0.25)}{350}} \\ n &= \frac{0.9604}{0.01 + \frac{0.9604}{350}} \\ n &= \frac{0.9604}{0.01 + 0.002744} \\ n &= \frac{0.9604}{0.012744} \\ n &\approx 75.352 \approx 76 \end{aligned}$$

According to the solution of this formula, the sample for the research project is seventy-six students belonging to the systems engineering program at the Universidad Francisco de Paula Santander Ocaña, who will be randomly selected from the total population.

2.2 Design of data collection instruments and data analysis techniques.

The survey will be applied to a specific number of students in the systems engineering program. A questionnaire of closed questions with categories and answer options that will be defined beforehand will be used as a collection element.

2.3 Data analysis

To analyze the information obtained and interpret the results, the following list of activities is presented:

- Inspection of the data obtained.
- Validate the reliability of the data collection instrument.
- Presentation of results.

2.4 Development

For the development of the application, we will use the agile methodology Scrum, which allows collaboration, flexibility, and incremental delivery of products. In Scrum, the project is divided into short cycles called "sprints," which last two to four weeks. Each sprint culminates in a functional increment of the software, which can be reviewed and adjusted as needed. The following is a brief description of how we will implement this methodology (see Fig. 1).

The first step to conducting this methodology is to define the Scrum team, then the requirements are gathered and documented, then the product backlog is created, which is the prioritized list of features, enhancements, and bug fixes to be implemented in the application.

Subsequently, the sprint planning is performed where the team selects the tasks to be developed in this sprint, and the sprint objective is stated, then the sprint execution is performed where the team is in the development of the tasks selected in the sprint planning, in this execution process, 15-minute meetings are held per day called "sprint meeting", 15-minute meetings are held daily called Daily where the team reports what they have done and how they are doing with their tasks.

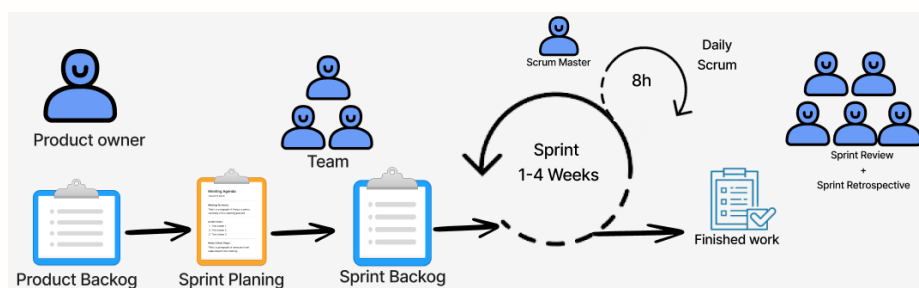


Fig. 1. Project development.

It continues with the Sprint retrospective, where at the end of the sprint, the team presents the incremental product (the part of the application developed during the sprint), finally, the Sprint retrospective is performed where the team reflects on the sprint process, discusses what went well, what can be improved and how they can make changes for the next sprint, this process is repeated until the application is complete and meets expectations.

3 Results

Hearing loss is one of the diseases that has become more important in recent years worldwide, so having tools that allow early diagnosis, and timely treatment is one of the objectives of global health organizations. In addition, it is important to know which side effects most frequently affect people with hearing loss, among which depression, anxiety, difficulty in communication, and deterioration of their daily quality of life stand out [8].

Now, to have a more realistic understanding of these effects that are currently affecting people with hearing loss problems, a measurement instrument (survey) was designed and applied to thirty people suffering from diagnosed and undiagnosed hearing difficulties. Each of the people suffering from hearing loss proceeded to select the frequency with which they presented some of the effects presented above or the different difficulties they face day by day due to hearing loss, using a 5-level Likert scale (Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree). With the application of this instrument, the effects that most affect hearing-impaired people were analyzed. As a result of the measurement instrument applied, the effects that most affect the hearing impaired were confirmed and identified. The survey included fifteen questions, each of

which pertains to or refers to one of the effects mentioned above, and below I will detail the questions that pertain to each effect.

Anxiety:

1. I experience anxiety or stress in situations where I cannot hear clearly.
2. "I avoid meetings or social activities because of my difficulty listening."
3. "I avoid activities in noisy spaces because I find it difficult to communicate."
4. Hearing loss affects my sense of security in everyday situations, such as crossing the street or identifying alerts.

Depression:

1. "I feel emotionally isolated because of my hearing problems."
2. "Hearing loss negatively affects my self-esteem."
3. "My relationships are affected by my hearing loss."
4. "People are often impatient or frustrated with me because of my hearing loss."

Communication problems:

1. "I am frustrated by the difficulties in listening to conversations."
2. "I have trouble participating in group conversations."
3. "I find it difficult to have phone conversations."
4. "I have difficulty understanding instructions or information at work."
5. "I feel that my communication skills affect my integration into my community or social environment."

Quality of life:

1. "My difficulty listening affects my ability to perform daily tasks."
2. "I feel like I don't have enough access to tools or services that can improve my hearing."

ITEMS - CANTIDAD DE PREGUNTAS																
ENCUESTADOS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	SUMA
E1	1	1	1	1	1	3	2	4	1	1	1	2	3	1	3	26
E2	5	5	1	2	2	2	2	4	4	2	2	2	2	2	2	39
E3	4	2	3	3	4	4	4	5	4	5	4	2	3	5	5	57
E4	4	4	5	4	3	5	4	3	4	4	4	4	4	5	3	60
E5	5	5	4	4	4	3	4	5	4	5	5	5	5	4	5	67
E6	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	60
E7	4	5	4	5	4	5	4	4	4	3	4	4	2	4	5	61
E8	4	4	4	4	3	4	4	4	4	4	4	5	5	4	5	62
E9	4	4	4	4	4	4	5	4	5	4	5	5	4	4	5	65
E10	4	4	4	3	3	3	3	4	4	4	3	4	2	4	5	54
E11	4	5	3	2	2	3	4	4	4	4	3	4	2	3	3	50
E12	4	4	5	3	4	3	4	3	4	3	4	4	4	4	5	58
E13	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	59
E14	3	3	3	4	4	4	3	3	4	4	4	4	3	4	4	54
E15	4	3	3	2	3	4	2	4	4	4	4	4	2	3	5	51
E16	5	4	2	3	3	4	4	4	4	3	4	4	4	4	4	56
E17	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	60
E18	4	5	4	5	3	4	5	5	4	5	4	4	3	3	3	61
E19	1	1	2	3	2	1	1	3	2	3	3	1	2	1	5	31
E20	5	2	5	3	4	4	1	3	4	1	3	4	1	5	5	50
E21	4	4	4	4	4	5	3	3	5	5	3	4	3	4	5	60
E22	4	4	4	4	4	4	3	3	3	3	3	4	3	4	4	54
E23	4	4	5	4	4	4	4	3	4	4	4	5	3	4	5	61
E24	4	4	4	4	4	5	4	3	4	4	4	3	3	4	4	58
E25	1	4	4	1	1	1	1	1	1	1	1	1	1	1	1	21
E26	4	4	2	2	2	2	2	4	4	2	2	3	3	4	4	44
E27	4	4	4	5	5	5	3	4	4	5	4	3	2	4	4	60
E28	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	75
E29	3	4	3	3	3	4	3	3	3	3	3	3	3	3	3	47
E30	4	5	4	4	3	4	3	3	3	5	3	3	2	3	2	51
VARIANZA	1,093	1,139	1,173	1,179	1,022	1,143	1,277	0,689	0,862	1,440	0,983	1,173	1,133	1,173	1,166	
SUMATORIA DE VARIANZAS	16,647															
VARIANZA DE LA SUMA DE LOS ITEMS	133,862															

Fig. 2. Questions.

In addition, to determine the reliability of the instrument, Cronbach's alpha was used. To implement Cronbach's alpha, considering the number of respondents and the number of questions in the instrument, and considering the 5-level Likert scale, the responses of each respondent were summed considering each of these levels. This

allowed us to know the variance and, with it, to obtain the sum of the variances and the variance of the sum of the items (see Fig. 2).

To determine the reliability of the instrument, we used Cronbach's alpha formula. Then, considering the range to define the reliability of the instrument and considering that the instrument had a reliability coefficient of 0.94, we can affirm that it presents excellent reliability.

3.1 Components considering the requirements posed by the clients/users.

Considering the design of the components, modules, and services with which the platform will be developed is of utmost importance since its operation and the logic applied to the concepts that are handled depend on it, we have worked with Angular considering the MVC ("Model-View-Controller") system architecture.

3.2 Software requirements

For this research project and to provide a solution to the hearing impaired through a multiplatform system, the following software requirements were considered for its development:

- Web accessibility with contrasting colors and quick access buttons.
- Simple and clear interfaces.
- Responsive development.
- Application of technical languages such as HTML5, CSS3, and TypeScript.
- Accessibility frameworks such as Angular 16 and Spring Boot.
- Storage of results.
- SSL for data and password encryption.
- Interactive hearing tests and their immediate results.

3.3 Architecture for web applications

The established architecture design is MVC (Model-View-Controller), considering that this architecture is the most suitable for Angular and Spring Boot, the programming languages used in the development of the research. In addition, it is due to its simple logic, since the model manages the data and business logic, the view is the part of the user interface that the client will see, and the controller is the component that will help manage the interaction between models and views. In addition, this architecture facilitates maintenance since each part of the code has a specific and defined responsibility. This allows us to adjust in a clear and specific way. This architecture also contributes to the scalability of our application, as it provides an ordered and clear structure (see Fig. 3).

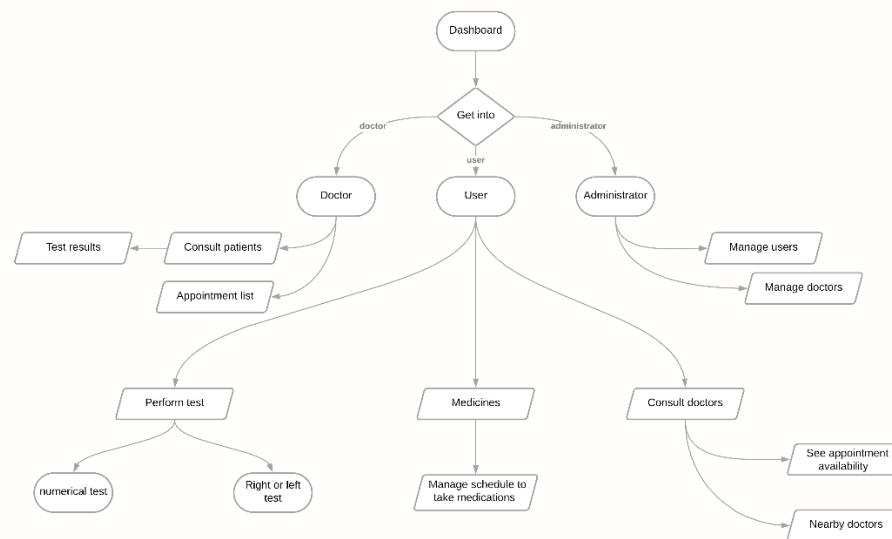


Fig. 3. Platform architecture.

3.4 Apply the SCRUM agile development methodology to oversee the project documentation process.

Currently, the use of the agile methodology SCRUM in the development of software projects is of great relevance, as it enables efficient management and helps to manage project uncertainty. SCRUM is an agile framework that fosters collaboration, flexibility, and continuous value generation, which makes it an ideal choice for software development [19–21].

- **Create the vision of the project.** This process was fundamental to establishing a clear objective aligned with the needs of the hearing impaired.
- **Identify the Scrum Master and stakeholders.** According to the SBOK™ Guide, identifying the Scrum Master and stakeholders is crucial to the success of the project. Magreth Sanguino assumed the role of Scrum Master, facilitating the process and ensuring compliance with SCRUM practices. The hearing impaired were identified as key stakeholders, given their role in providing feedback on product requirements and functionality.
- **The SCRUM Team.** According to the SBOK™ Guide, a SCRUM team should be composed of professionals with cross-functional skills who work autonomously. In this project, the team was composed of two members: José Quintero and Diany García. This structure facilitated direct communication and greater cohesion, essential aspects for the success of the project.

For the development of our platform, several technologies were implemented for both the front-end and back-end. On the front-end, HTML5 was used to structure the basic components, which ensures proper visualization for users when accessing the platform and facilitates a well-structured and organized code. CSS3 was also used to define the system-specific layouts and styles, and TypeScript was used to manage the functions and processes required to display the information obtained from the back end. On the other hand, the back end was developed with Spring Boot to implement security, manage business logic, and perform queries to the database, developed in PostgreSQL (see Fig. 4).

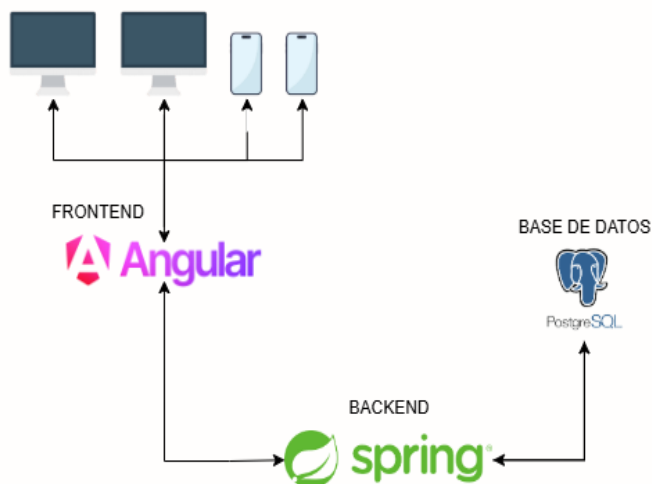


Fig. 4. Frontend and backend relationship.

3.5 Schema for storing medical information in relational databases such as PostgreSQL.

PostgreSQL was chosen as the database for this project because of its reliability, scalability, and robustness in handling complex relationships, which guarantees the integrity and security of critical data in a healthcare platform. In addition, its open-source code allowed us to reduce costs without compromising performance (see Fig. 5).

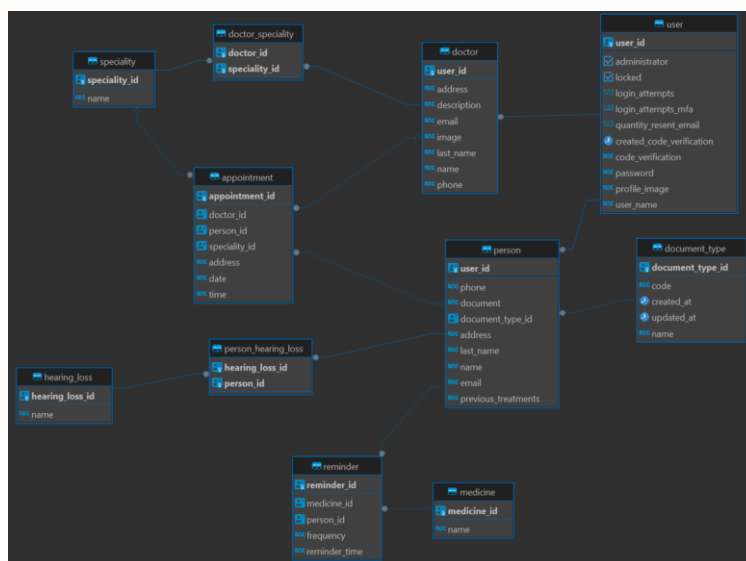


Fig. 5. Database schematic.

3.6 Configuration of development environments

To ensure that the development environments are configured correctly with the corresponding versions of Angular, Spring Boot, and PostgreSQL, we perform the following steps:

- Install Node.js or otherwise check the versions, as this is the fundamental basis for Angular projects and the installation of packages with npm.
- Angular CLI 16 was installed globally with the help of npm, which allowed us to visualize and manage the project with its dependencies (Angular allowed to run projects in previous versions if you have a new one).
- Initialize the project and install the dependencies, verify that the package.json file is well configured since it contains all our dependencies with their versions for the project.
- Installation of dependencies
- Look at the configuration of the package. json
- Configure the Spring Boot environment by installing the correct version of the JDK, which in this case is version 17, and select the essential dependencies such as the JDBC connector for PostgreSQL.
- Configure the build.gradle file for the above dependencies, checking versions and libraries.
- The installation, creation, and configuration of our database with PostgreSQL is performed.
- Perform the respective integrations of the Front with the Back and the Back with the database and also perform connection tests between each one to ensure that the development environment is already well configured.
 - Front-End configuration to guide the Back End.
 - Configuration of the Back end to receive requests from the Front-End
 - Connection to the database from the Back End

3.7 Develop modules in the system.

The modules developed for the platform are User, Aut., Doctor, and Administrator. The Aut. The module includes functions such as login, registration, and password recovery for both the user and the doctors. To log in, only the previously created user and password are required. To register, a form must be completed with basic personal data such as name, surname, e-mail, telephone number, and type of document, among others. Finally, in this module, you can recover your password, which is obtained through the registered e-mail address. The Doctor module includes an administrative panel where the doctor can view his patients, consult data, modify his data, and modify his schedule. The interface is simple and convenient for doctors. The administrative module, on the other hand, has a series of tables where doctors can be deleted or added to the platform, as well as activate or deactivate user accounts. The administrator can only see the user's name and status and will not have access to other information to protect patient data.

3.8 Software performance testing for quality assurance

Perform tests on the different modules, such as tests, verify that they are executed correctly, that the corresponding errors or information messages are displayed depending on the result of each one, and that the information is sent or saved correctly for that user.

3.9 Validation of the accessibility and usability of the application.

Considering that the sample previously stipulated for the project was seventy-six students of the systems engineering program of the Universidad Francisco de Paula Santander Ocaña, the survey was conducted among this number of students, from which the following results were obtained. Regarding the age of the students, it was found that the average age is between 16 and 30 years old. Considering the activity to be performed, questions were created to evaluate the accessibility and usability of the Sound Good platform developed for this research work. According to these questions, the following results were obtained. From the results presented above, it can be concluded that the platform meets the necessary accessibility and usability, and the vast majority of respondents agreed with this [22–25].

In addition to the survey conducted, the accessibility of the website was validated with Lighthouse, a Google tool that provides a detailed report on accessibility, best practices, and SEO. The results are expressed on a scale from 0% to 100%, which made it possible to identify errors or aspects to be improved to achieve complete optimization. Lighthouse is an automated open-source tool that, in addition to assessing accessibility, analyzes website performance, determines loading speed and efficient use of resources, and reviews compliance with modern web development standards. In addition, the tool allowed saving the test results in different formats, such as JSON, HTML, or PDF, thus facilitating the documentation and monitoring of the implemented improvements. Taking into account the results obtained from the platform with the Lighthouse tool, a survey composed of 15 questions was developed to evaluate the accessibility and usability of the platform, for this purpose and taking into account the sample previously chosen for the project, the survey was applied to 76 currently active students of the systems engineering program at the Universidad Francisco de Paula Santander Ocaña [26], from which the following results were obtained, see Fig. 6.

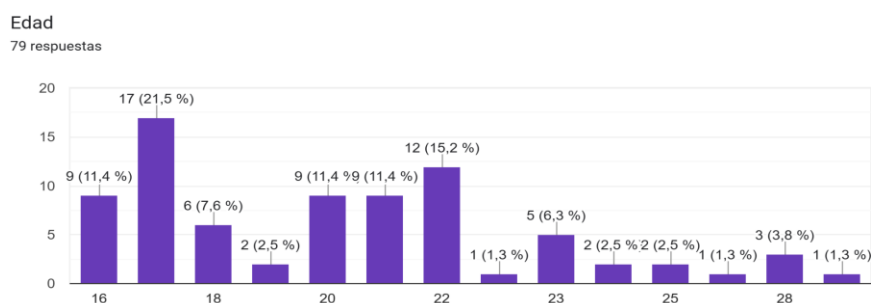


Fig. 6. Usability.

3.10 Feedback from hearing health professionals

The purpose of this activity was to have feedback with healthcare professionals about the developed web platform, in order to evaluate usability and functionality from a clinical point of view and thus identify strengths and improvements for our platform.

For this purpose, two healthcare specialists were consulted, and interviews and live tests were conducted with the platform via video call, the main objective being that the healthcare specialists could identify the ease of use, and verify the interface design, key functionalities, and possible improvements for the platform.

The results obtained from this feedback, where specialists agreed, indicate that the platform presents an intuitive interface, with clear navigation and easy to understand. It was also identified that the content of the platform concerning relevant information is useful and well structured, and professionals see enormous potential in the platform with a high value of use for monitoring and guidance to patients. Also, professionals consider the platform to have adequate contrast, and the speed of data loading is also considered to be correct. The healthcare professionals with whom this feedback was conducted concluded that the platform is suitable for use in the healthcare sector, where it is easy navigation, relevant content, and high application potential have been highlighted. They consider the platform in its current state to be a useful tool for specialists and patients (see Fig. 7).

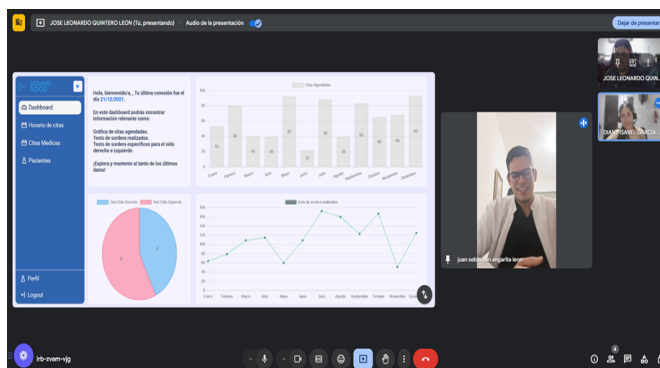


Fig. 7. Image used with permission of the people shown.

4 Conclusions

This paper addressed aspects of hearing loss, the prediagnosis of hearing loss and the effects it can have, as well as the importance of implementing technologies for early and accurate prediagnosis of hearing loss.

It should also be noted that, although there are programs used for the prediagnosis of hearing loss, many of them are outdated, resulting in the obsolescence of technological detection tools.

Finally, we propose the development of a multiplatform system, where multiplatform means that it can be run in any operating system or environment, to help us in the pre-diagnosis of hearing loss, where people with disabilities can perform hearing tests from the platform and thus provide an early pre-diagnosis, since many of these people have difficulty in approaching a hospital center for studies, also an additional platform is that the patient can schedule medical appointments from the same, with specialists who are in the database.

The system still has room for improvement. There are plans to optimize the database and connect it with the databases of the EPS or health centers, which will strengthen the process of scheduling medical appointments. In addition, the inclusion of new hearing tests and the generation of official medical certificates on the state of hearing quality is an improvement.

To carry out the evaluation process of our project we will use an instrument that will allow us to obtain quantitative and qualitative data, which will provide us with a more global vision of the results obtained, focusing on the fulfillment of the objectives and areas for improvement that can be identified, preserving the long-term sustainability and quality of our proposal.

User experience is key to the project. To achieve this, usability tests will be included to help evaluate ease of use, intuitive navigation, and user satisfaction. The feedback from these tests will be fundamental for the continuous update and improvement of our application.

After the research was conducted, it was found that the term "hearing impairment" does not have an exact date of origin, as it has evolved simultaneously with the understanding of hearing impairment and the terminology used in the medical field. In many cases, the evolution of the term and the concept is linked to the development of audiology as a discipline and to technological progress in the field of hearing impairment and its intervention approaches.

Some recommendations are as follows, keeping in mind that other students may want to continue with this research project idea:

- Incorporate more accessibility tools into the platform, such as voice recognition and real-time captioning, as well as the implementation of chatbots.
- Make strategic alliances with health entities and organizations that support people with hearing loss for the implementation of the platform.

- Evaluate the possibility of being able to connect the platform with digital audiometers.
- Implement the use of artificial intelligence to improve the analysis of hearing test results.

References

1. Baker A (2003) Psychiatric problems in people with hearing loss. *Rev Afr Polit Econ* 30:479–510
2. Baker J (2012) The Technology–Organization–Environment Framework. pp 231–245
3. Bassiouni M, Häußler SM, Ketterer MC, Szczepek AJ, Vater J, Hildebrandt L, Gröschel M, Olze H (2023) Prospektive Vergleichsanalyse von CI-Patienten mit einseitiger Taubheit und asymmetrischem Hörverlust hinsichtlich der gesundheitsbezogenen Lebensqualität, Tinnitusbelastung und psychischen Komorbiditäten. *HNO* 71:494–503
4. Brewster KK, Ciarleglio A, Brown PJ, Chen C, Kim H-O, Roose SP, Golub JS, Rutherford BR (2018) Age-Related Hearing Loss and Its Association with Depression in Later Life. *The American Journal of Geriatric Psychiatry* 26:788–796
5. Sharma A, Rutherford BR, Lin F, Golub JS, Brewster KK (2019) AGE-RELATED HEARING LOSS AS A RISK FACTOR FOR LATE LIFE DEPRESSION AND COGNITIVE DECLINE. *The American Journal of Geriatric Psychiatry* 27:S43–S44
6. Lawrence BJ, Alexander E, Jayakody DMP (2021) Hearing Loss and Depression in Older Adults: A Small Association. *Am J Med* 134:e76–e77
7. Spanoudakis G, Katrakazas P, Koutsouris D, Kikidis D, Bibas A, Pontopidan NH (2017) Public Health Policy for Management of Hearing Impairments Based on Big Data Analytics: EVOTION at Genesis. In: 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE, pp 525–530
8. Lu Z, Yu D, Wang L, Fu P (2024) Association between depression status and hearing loss among older adults: The role of outdoor activity engagement. *J Affect Disord* 345:404–409
9. Lersilp T, Lersilp S (2019) Use of Information Technology for Communication and Learning in Secondary School Students with a Hearing Disability. *Educ Sci (Basel)* 9:57
10. Kyburz Philip, Gfeller Sascha, Bürkle Thomas, Denecke Kerstin (2019) Exchanging Appointment Data Among Healthcare Institutions. <https://doi.org/10.3233/978-1-61499-971-3-33>
11. Hung Veronica, Fung Vicky, Lau Erica, Lau Stephen, Lam Dick, Pow Mavis (2019) Realizing the Benefits of Managing Health Appointments via Mobile Application: Start of the Journey. <https://doi.org/10.3233/SHTI190597>
12. Heitmann N, Rosner T, Chakraborty S (2021) Mass-deployable Smartphone-based Objective Hearing Screening with Otoacoustic Emissions. In: Proceedings of the 2021 International Conference on Multimodal Interaction. ACM, New York, NY, USA, pp 653–661
13. Wongsirichot T, Elz N, Suwanmanee N, Hmad-a-dam H, Plansangket S, Sukpisit S (2017) Hearing Performance Testing Application. In: Proceedings of the 1st International Conference on Medical and Health Informatics 2017. ACM, New York, NY, USA, pp 7–11
14. Hernández Montero G (2021) Historia de la audiolología. *Revista Cubana de Otorrinolaringología y Cirugía de Cabeza y Cuello* 5:
15. Hernández Sampieri R, Fernández Collado C, Baptista Lucio P (1997) Metodología de la investigación.
16. Hernández, R., Fernández, C. y Baptista P (2014) Libro Metodología de la Investigación 6ta edición SAMPIERI (PDF) | Metodologiaacs.
17. Danel O (2015) Metodología de la investigación. Técnicas e instrumentos. <https://doi.org/10.13140>
18. Rivas L (2016) Elaboración de Tesis (Estructura y metodología). *Revista Trillas* 1:1–384
19. Cervone HF (2011) Understanding agile project management methods using Scrum. *OCLC Systems & Services: International digital library perspectives* 27:18–22
20. Singh M (2008) U-SCRUM: An Agile Methodology for Promoting Usability. In: Agile 2008 Conference. IEEE, pp 555–560
21. Arango EC, Loaiza OL (2021) SCRUM Framework Extended with Clean Architecture Practices for Software Maintainability. pp 667–681
22. Romero-Riaño E, Galeano-Barrera C, Guerrero CD, Martínez-Toro M, Rico-Bautista D (2022) IoT applied to irrigation systems in agriculture: A usability analysis. *Revista Colombiana de Computación* 23:44–52
23. Jacob Nielsen (1994) Usability Engineering.
24. Rocca JI, Solari M (2018) Análisis de Procesos de Software Orientados a la Usabilidad.
25. Granollers T (2018) Usability Evaluation with Heuristics. New Proposal from Integrating Two Trusted Sources 2 Combining Common Heuristic Sets. 1–16
26. Areniz-Arevalo Y, Barrientos-Avendaño E, Coronel-Rojas L, Cuesta-Quintero F, Bayona-Ibañez E, Rico-Bautista D (2020) Plataforma Moodle como apoyo a la presencialidad: Experiencia de uso en la Universidad Francisco de Paula Santander Ocaña en tiempo de Covid-19. *Revista de la Asociación Interacción Persona Ordenador (AIPO)* 1:74–82

Narrativa Semántica y Características Prosódicas para Audiodescripción

Cristian Plazas¹ [0009-0008-2633-9607], Juan Sebastian Cardenas¹ [0009-0005-5583-7427],
Christian Quintero¹ [0000-0001-5763-2373]
Universidad Militar Nueva Granada, Bogotá, Colombia
cristian.plazas534@gmail.com, {est.juan.cardenas8,
christian.quintero}@unimilitar.edu.co

Resumen. La audiodescripción es fundamental para la accesibilidad de personas con discapacidad visual, ya que les permite acceder a contenido visual mediante narraciones descriptivas. Sin embargo, los procesos actuales de creación de contenido presentan limitaciones en la generación de descripciones semánticamente precisas y en la incorporación de prosodia natural, lo que afecta la calidad y comprensión de las narraciones. Objetivo: Proponer un esquema de trabajo basado en mecanismos de atención que genere descripciones semánticas y prosódicas para mejorar la calidad de las audiodescripciones. Métodos: Se diseñó un esquema que extrae información semántica de fotogramas clave, analizados mediante un conjunto de datos etiquetado con mecanismos de atención para obtener características semánticas y prosódicas. El esquema incluyó un procesamiento de datos mediante un modelo generativo. La evaluación se realizó comparando los resultados con otros modelos de descripción y mediante un análisis mixto basado en la similitud de las descripciones. Resultados: El esquema propuesto mostró mejoras significativas en la precisión semántica en comparación con los modelos existentes. Además, generó una guía prosódica que permite realizar narraciones con diferentes intenciones emocionales, adaptadas al contexto de la escena. Conclusiones: Este trabajo demuestra que la integración de mecanismos de atención y generación de prosodia puede transformar la calidad de la audiodescripción automática o tradicional, ofreciendo narraciones más naturales. Se identifican nuevas oportunidades para optimizar el esquema y ampliar su aplicabilidad en evaluaciones con usuarios.

Palabras Claves: Audiodescripción, Descripción Automática, Mecanismos de Atención, Prosodia.

1 Introducción

El acceso equitativo al contenido audiovisual continúa siendo un desafío muy importante, especialmente para personas con discapacidades visuales, debido a la limitada disponibilidad de descripciones precisas y adaptadas que faciliten una comprensión plena de estos elementos visuales y narrativos. En este contexto, la audiodescripción automática basada en modelos generativos que han emergido como una solución para mejorar la accesibilidad digital generando automáticamente descripciones textuales que representen de manera natural el contenido visual de los videos, permitiendo así una accesibilidad digital más inclusiva.

Aunque los modelos actuales de descripción automática han avanzado considerablemente, principalmente gracias al desarrollo de las redes neuronales recurrentes, aún enfrentan limitaciones en la generación de descripciones detalladas. La mayoría de los casos se enfocan en imágenes [1], mientras que el progreso en videos es mucho más limitado. Además, las narrativas semánticas requieren mejoras importantes para satisfacer plenamente las necesidades de los usuarios con discapacidades visuales.

La audiodescripción automática ha evolucionado significativamente en los últimos años, impulsada por avances en visión por computadora y procesamiento de lenguaje natural. Esta sección analiza los trabajos más relevantes en tres áreas fundamentales: narrativa semántica para descripción de videos, mecanismos de atención para mejorar la precisión descriptiva, y características prosódicas en narración automática.

La Narrativa Semántica (NS) se ha convertido en una técnica clave, ya que permite identificar y describir múltiples eventos dentro de un video sin segmentación manual previa. Este enfoque utiliza Redes Neuronales Convolucionales (RNC) y modelos basados en transformadores para generar descripciones detalladas y coherentes, capturando tanto el contexto temporal como espacial de los eventos [2, 3]. Además, técnicas

avanzadas como la fusión multimodal de datos han permitido integrar información visual, auditiva y textual, mejorando significativamente la contextualización y riqueza de las descripciones [4].

Los mecanismos de atención han surgido como una herramienta fundamental para mejorar la precisión y relevancia de las descripciones generadas automáticamente. Estos mecanismos permiten que los modelos enfoquen selectivamente su procesamiento en las regiones o momentos más informativos de una secuencia de imágenes, "enfocándose precisamente en las partes importantes mientras suprimen otra información inútil". Al integrar mecanismos avanzados como el módulo de atención coordinada (CA) en arquitecturas como los transformadores, el sistema puede determinar dinámicamente los pesos de las características del video e identificar qué partes del contenido requieren mayor énfasis, incorporando información posicional crucial que los enfoques tradicionales ignoraban. Esta capacidad de focalización ha demostrado ser especialmente útil en escenarios complejos donde múltiples eventos ocurren simultáneamente, resultando en audiodescripciones más precisas, coherentes y significativas para los usuarios [5].

Las características prosódicas han surgido como una estructura fundamental dentro del estudio fonológico del lenguaje, debido a su estrecha relación con otras estructuras lingüísticas como la sintaxis, el discurso y la organización de la información. Estas características permiten que el habla transmita no solo contenido verbal, sino también matices rítmicos y expresivos que facilitan la comprensión y la interacción comunicativa. Las características prosódicas pueden dividirse en tres aspectos principales: entonación, distribución temporal y acento, implementadas a través de rasgos suprasegmentales. Dichos rasgos incluyen la altura tonal (el rango de tono en la frecuencia fundamental de la voz), la intensidad (energía sonora o amplitud) y las características temporales. Lingüísticamente, el término "prosodia" hace referencia a estas propiedades suprasegmentales del habla, incluyendo la altura tonal, la intensidad y las variaciones en la velocidad del habla y las pausas [6].

Este trabajo propone un esquema basado en mecanismos de atención para la generación de descripciones semánticas y prosódicas, estableciendo una guía para la narración en audiodescripciones. El objetivo principal es no solo garantizar descripciones detalladas y coherentes con el contenido visual, sino también ofrecer una experiencia auditiva emocionalmente enriquecida, adaptada a las necesidades específicas del público objetivo. Por medio de la integración de elementos semánticos y prosódicos, este enfoque busca describir a través de narrativas e intenciones emocionales de las audiodescripciones, fomentando la accesibilidad digital inclusiva y promoviendo un entorno más equitativo en el acceso a la información y el entretenimiento.

2 Trabajos relacionados

La literatura señala diversos esfuerzos orientados a perfeccionar la descripción automática en entornos audiovisuales. Entre ellos destacan enfoques de NS, un término utilizado en este trabajo que se orienta a técnicas como se denomina en inglés -Dense Video Captioning-. Además, se abordan los Mecanismos de Atención (MA) y el Procesamiento de Lenguaje Natural (PLN), los cuales han demostrado ser fundamentales para mejorar la calidad y precisión de las descripciones generadas.

Con respecto a las NS, el trabajo Shen et al [4] combina características visuales y auditivas para generar descripciones detalladas y contextualmente enriquecidas, mejorando la accesibilidad y comprensión del contenido audiovisual. Así mismo, la generación de descripciones para múltiples eventos en videos propone una arquitectura de redes neuronales recurrentes densamente conectadas, lo que permite una mayor retención de contexto y precisión en las descripciones generadas [7].

Por otro lado, el modelo AutoAD-Zero [8] introduce un sistema de generación de audiodescripciones sin necesidad de entrenamiento previo, optimizando la creación de narrativas comprensibles mediante un proceso de dos etapas que incluye la descripción detallada y su síntesis. Además, la implementación de técnicas de fusión multimodal, como las utilizadas en MemCap [9], muestra ser efectiva para integrar información visual, textual y auditiva, generando descripciones estilizadas y coherentes.

En consecuencia, los proyectos no solo han mejorado la precisión y naturalidad de las descripciones generadas, sino que también han abierto nuevas oportunidades para la accesibilidad digital. Sin embargo, no se

evidenciaron trabajos en la descripción dirigidos en prosodia, por lo tanto, es una oportunidad de trabajar este campo que permite un enfoque hacia a la experiencia del usuario con discapacidad visual.

Los MA han demostrado ser esenciales para optimizar el procesamiento de información, permitiendo que los modelos se enfoquen dinámicamente en las partes más relevantes de las entradas visuales del video. Esto mejora la precisión y coherencia de las descripciones generadas, como se evidencia en proyectos que combinan las Redes Neuronales Recurrentes con la arquitectura de Memoria a Largo y Corto Plazo y MA avanzados [4, 8, 9]. Asimismo, el procesamiento de lenguaje natural ha sido clave para garantizar que las descripciones sean no solo precisas, sino también fluidas y comprensibles, adaptándose a las necesidades de los usuarios finales [11].

A pesar de los avances significativos en NS y MA, existe una notable ausencia de investigaciones centradas específicamente en la incorporación de características prosódicas en la audiodescripción. Esta brecha representa una oportunidad importante para mejorar la experiencia del usuario con discapacidad visual, por medio del modelado prosódico y sus características de como intensidad, tono, ritmo y entonación es fundamental para transmitir emociones y matices que complementan la información semántica.

Los trabajos existentes se han centrado principalmente en la precisión semántica de las descripciones, pero han prestado poca atención a cómo estas descripciones deben ser narradas para maximizar su efectividad comunicativa. Esta limitación es particularmente relevante considerando que la audiodescripción no solo debe proporcionar información precisa sobre el contenido visual, sino también transmitir adecuadamente la carga emocional y el contexto narrativo de las escenas descritas.

Por lo tanto, este trabajo busca abordar esta brecha mediante la integración de MA para la generación de NS. Además, se complementa con características prosódicas que brinda una correspondencia con una intencionalidad emocional del contenido visual, mejorando así la experiencia auditiva para personas con discapacidad visual.

3 Esquema propuesto

El esquema propuesto para la generación de audiodescripciones automáticas combina técnicas de análisis de imágenes, extracción semántica y características prosódicas, con el objetivo de mejorar la accesibilidad y enriquecer la experiencia narrativa para personas con discapacidades visuales. Este enfoque se estructura en varias etapas interconectadas, como se detalla a continuación:

- **Diarización Inversa y Segmentación del Video.** El proceso comienza con los resultados de una diarización, en donde se identifican los segmentos del video donde hay presencia de hablantes y a partir de estos podemos obtener los segmentos sin hablantes. Esto permite una segmentación independiente de la escena y estructurada del contenido, favoreciendo el análisis posterior.
- **Separación de Escenas.** Una vez obtenidos estos segmentos, se realiza un análisis para verificar cambios de escenas, seleccionando únicamente aquellas que duran más de dos segundos, lo que garantiza que haya suficiente tiempo para insertar una descripción narrativa coherente. Esta verificación temporal es crucial para que la audiodescripción sea efectiva, ya que escenas demasiado breves no permitirían una narración adecuada.
- **Extracción Semántica de la Escena.** Se realiza una extracción semántica de fotogramas clave utilizando un modelo de aprendizaje profundo Preentrenamiento Contrastivo de Lenguaje e Imágenes (CLIP, por sus siglas en inglés) [12], la cual aprende eficientemente conceptos visuales a partir de la supervisión del lenguaje natural que se entrena juntamente con un codificador de imágenes y un codificador de texto para predecir correctamente las combinaciones entre pares. Finalmente, realiza un proceso de selección de fotogramas claves los cuales predominan en la escena.
- **Generación de Narración Semántica y Características Prosódicas.** Se generan descripciones detalladas de cada fotograma clave utilizando inteligencia artificial generativa, basada en mecanismos de atención que permiten a los modelos enfocarse en las partes más relevantes de una entrada de datos, asignando pesos específicos a cada elemento según su importancia. Este enfoque, denominado narrativa semántica, se centra en identificar y describir objetos, acciones y contextos relevantes, lo que

permite generar descripciones precisas y coherentes que capturan la esencia de cada segmento del video. De forma simultánea, se realiza una extracción de características prosódicas, como la intensidad, el tono, el ritmo y la entonación, asociadas a una intención emocional. Este componente prosódico enriquece la experiencia auditiva, asegurando que la narración sea clara, coherente y emocionalmente alineada con la escena descrita.

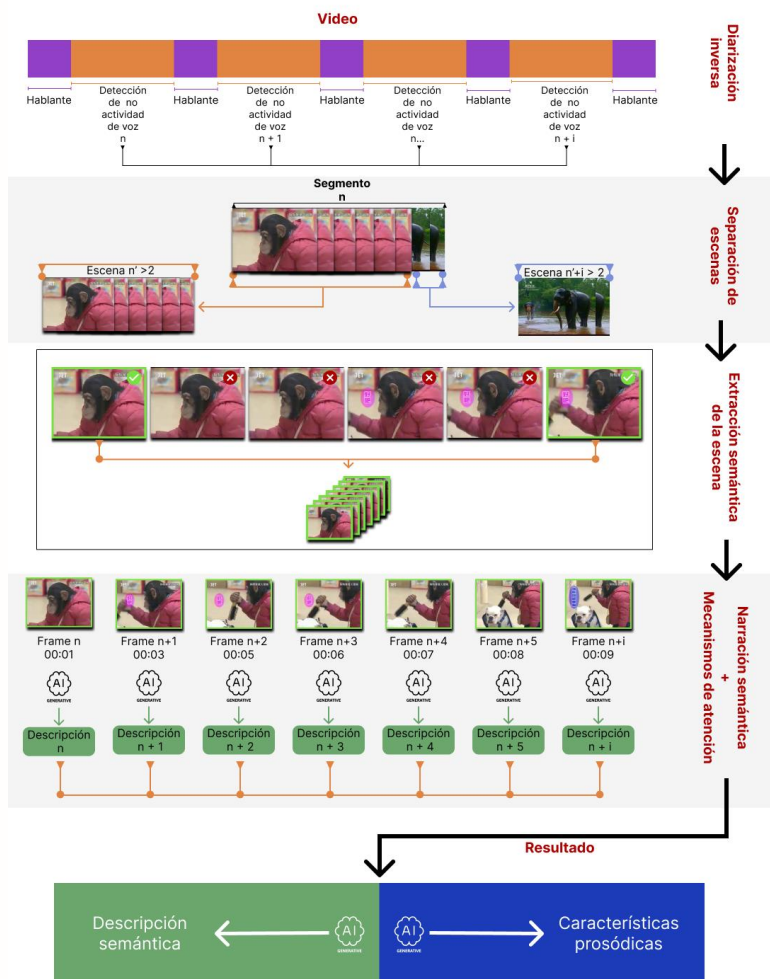


Fig. 1. Diagrama de flujo del esquema propuesto para la descripción narrativa semántica automática.

4 Metodología

Para el desarrollo del modelo propuesto, se utilizó el conjunto de datos Microsoft Research Video Description Corpus (MSVD)[13], cuya estructura incluye videos cortos acompañados de múltiples descripciones generadas manualmente. Este diseño permite evaluar directamente la calidad narrativa y calcular métricas objetivas relacionadas con la similitud semántica. Según Islam et al [14], el conjunto consta de 1,800 fragmentos de video con 120,000 anotaciones humanas en múltiples idiomas. La mayoría de los grupos de investigación han dividido este conjunto de datos en particiones de entrenamiento, validación y prueba, con 1,200, 100 y 670 videos, respectivamente. En todos los videos, el audio está silenciado, y se han omitido aquellos que contienen

subtítulos o texto superpuesto. Por lo general, la duración de cada video oscila entre 10 y 25 segundos, y la mayoría representa una sola actividad. Estos videos se utilizan en aplicaciones como asistencia para personas con discapacidad visual, interacción humano-robot, sistemas de videovigilancia y comprensión de escenas [1].

El experimento se diseñó para evaluar la capacidad del sistema de audiodescripción automática en la generación de narrativas claras y accesibles. Para ello, los videos fueron segmentados en fotogramas clave mediante el modelo CLIP. La segmentación fue adaptativa, ajustándose dinámicamente a la duración del video y su tasa de Fotogramas Por Segundo (FPS). A partir de los fotogramas extraídos, se generaron descripciones automáticas usando un modelo de inteligencia artificial generativa. Se estableció un límite de caracteres proporcional a la duración del video (30 caracteres por segundo). Adicionalmente, se incorporaron mecanismos de atención avanzados para mejorar la coherencia y el nivel de detalle en la descripción de cada escena.

Así mismo, se elaboró un análisis de prosodia, ajustando el tono, ritmo e intensidad de la narración con base en las emociones predominantes de cada video. A través del procesamiento de los fotogramas clave y las descripciones generadas, se identificaron emociones principales y secundarias, asignando porcentajes de relevancia para reflejar la intencionalidad emocional de la escena. Este paso permitió que la audiodescripción no solo fuera precisa y clara, sino también coherente con la carga emocional del contenido visual, mejorando así la experiencia de accesibilidad.

Por otra parte, la generación de prosodia se realizó a partir del siguiente prompt especializado:

- "Analiza las siguientes imágenes de frames del video y el resumen para determinar las emociones presentes: Resumen del video: \{combined_text\} [Frame Images: \{frames\}] Proporciona una única prosodia general para todo el video en el siguiente formato, basado en la tabla de referencia: Formato de salida (Intensidad: [texto según tabla] Tono: [texto según tabla] Duración: [texto según tabla] Ritmo: [texto según tabla].
- No des recomendaciones, interpretaciones adicionales ni ajustes sugeridos. Solo responde siguiendo el formato indicado para cada emoción principal detectada.
- Referencia prosódica:
- -Prosodia 1 (Intensidad: Baja; Tono: Bajo, estable, sin grandes variaciones; Duración: Larga, pausas relajadas; Ritmo: Lento, regular, pausado)
- -Prosodia 2 (Intensidad: Alta; Tono: Alto, con variaciones amplias; Duración: Corta, con exclamaciones; Ritmo: Rápido, energético, dinámico)
- -Prosodia 3 (Intensidad: Baja; Tono: Medio, cálido, poca variabilidad; Duración: Moderada; Ritmo: Regular, fluido, sin cambios bruscos)
- -Prosodia 4 (Intensidad: Alta; Tono: Alto, con variaciones ascendentes; Duración: Moderada; Ritmo: Enfático, con cierto realce, pero no brusco)
- -Prosodia 5 (Intensidad: Baja; Tono: Ligero ascenso, posible temblor o quiebre; Duración: Moderada, con pausas vacilantes; Ritmo: Irregular, inseguro, a veces acelerado)
- -Prosodia 6 (Intensidad: Alta; Tono: Alto, con quiebres marcados; Duración: Corta, entrecortada; Ritmo: Muy acelerado o brusco, pausas de jadeo)
- -Prosodia 7 (Intensidad: Muy baja e inconstante; Tono: Variable, a veces errático; Duración: Variable (puede ser corta o larga); Ritmo: Inconsistente, con cambios inesperados)
- -Prosodia 8 (Intensidad: Alta; Tono: Alto y ascendente al inicio, luego variable; Duración: Corta o moderada, con pausas breves; Ritmo: Brusco e irregular, marcando sorpresa)
- -Prosodia 9 (Intensidad: Baja; Tono: Bajo, levemente descendente; Duración: Larga, con pausas reflexivas; Ritmo: Lento, reflexivo, casi meditativo)
- -Prosodia 10 (Intensidad: Alta; Tono: Bajo, descendente, monótono; Duración: Larga, con pausas prolongadas; Ritmo: Muy lento, pesado, con silencios marcados)
- -Prosodia 11 (Intensidad: Baja; Tono: Bajo, con poca variación; Duración: Moderada-larga, habla sostenida; Ritmo: Lento, plano, sin cambios de intensidad)
- -Prosodia 12 (Intensidad: Alta; Tono: Alto, seco o cortante; Duración: Moderada, a veces con pausas bruscas; Ritmo: Cortante, agresivo, enfatizando el rechazo)
- -Prosodia 13 (Intensidad: Baja; Tono: Medio, con pequeños picos de tensión; Duración: Moderada, frases más cortas; Ritmo: Un tanto acelerado, interrumpido, con quiebres)

- -Prosodia 14 (Intensidad: Alta; Tono: Alto, agresivo, con cambios bruscos; Duración: Moderada o corta, intercalada con exclamaciones; Ritmo: Rápido, explosivo, con picos intensos)
- -Prosodia 15 (Intensidad: Baja; Tono: Alto, ligeramente ascendente; Duración: Variable; Ritmo: Fluido, ágil, mostrando atención sostenida)
- -Prosodia 16 (Intensidad: Alta; Tono: Alto, a veces tenso; Duración: Moderada, con pausas; Ritmo: Ágil, algo acelerado, atento a cambios)"

Finalmente, se realizó una evaluación comparativa de ocho modelos de inteligencia artificial generativa para la descripción de videos, utilizando como referencia una descripción oficial. Los modelos evaluados fueron SpaceTimeGPT [15], VideoChat-Flash [16], Video-LLaVA [17], Tarsier [18], ShareGPT4Video [19], Lc-Video-Description [20], Video-LLaMA2 [21] y nuestro modelo. La evaluación incluyó 10 videos seleccionados con descripciones en español que contaban con una única anotación, asegurando una diversidad de escenas y contextos visuales para realizar un análisis integral del desempeño de los modelos.

5 Resultados

El proceso de evaluación se basó en tres criterios principales: precisión, nivel de detalle y claridad y comprensión. La precisión midió la capacidad del modelo para identificar correctamente personajes, objetos y eventos dentro del video. Esto incluyó la identificación clara de los sujetos, sus características distintivas, las acciones que realizan y la contextualización adecuada del entorno y los escenarios. El nivel de detalle evaluó la inclusión de información relevante sin redundancias, mientras que la claridad y comprensión verificó la coherencia narrativa de las descripciones generadas.

Para la evaluación de la prosodia en la descripción de videos, se utilizó un enfoque basado en inteligencia artificial generativa, donde se analizaron los fotogramas clave de cada video junto con la descripción generada por nuestro modelo. A través de este análisis, se determinaron las características prosódicas más adecuadas para la narración, considerando emociones principales y complementarias, intensidad, tono, duración y ritmo.

La Figura 2 muestra una evaluación comparativa entre varios modelos de descripción de video en términos de tres criterios clave: precisión, nivel de detalle y claridad y comprensión. En el experimento, se analizaron 10 videos del conjunto de datos MSVD, con el objetivo de evaluar cómo diferentes modelos generan descripciones automáticas de los eventos visuales y cómo estas se alinean con las expectativas de los usuarios. Se procesaron la muestra de videos mediante los diferentes modelos de descripción automática seleccionados.

Tabla 1. Características prosódicas en la descripción de videos.

VideoID	Intensidad	Tono	Duración	Ritmo
zTn-nNj5Bng	Baja	Medio, cálido, poca variabilidad	Moderada	Regular, fluido, sin cambios bruscos
zpgW7m7_LZw	Baja	Bajo, estable, sin variaciones	Larga, pausas relajadas	Lento, regular, pausado
ZdlG8fjGJlo	Alta	Alto, con variaciones ascendentes	Moderada	Enfático, con cierto realce, pero no brusco
za-9mBZyNfQ	Alta	Alto, con variaciones amplias	Corta, con exclamaciones	Rápido, enérgico, dinámico
Z3XarHZ3HpY	Alta	Alto, a veces tenso	Moderada, con pausas	Ágil, algo acelerado, atento a cambios
z0Si1XxMibg	Baja	Bajo, estable, sin grandes variaciones	Larga, pausas relajadas	Lento, regular, pausado
YersIyzsOpc	Baja	Bajo, monótono o levemente descendente	Larga, con pausas reflexivas	Lento, reflexivo, casi meditativo
YEmS8ZQIhTA	Alta	Alto, con variaciones amplias	Corta, con exclamaciones	Rápido, enérgico, dinámico
yAD_TS5L2d4	Baja	Alto, ligeramente ascendente	Variable	Fluido, ágil, mostrando atención sostenida

XZ0Ch_CIBfo

Alta

Alto, con variaciones
amplias

Moderada-corta, con
exclamaciones

Rápido, enérgico,
dinámico

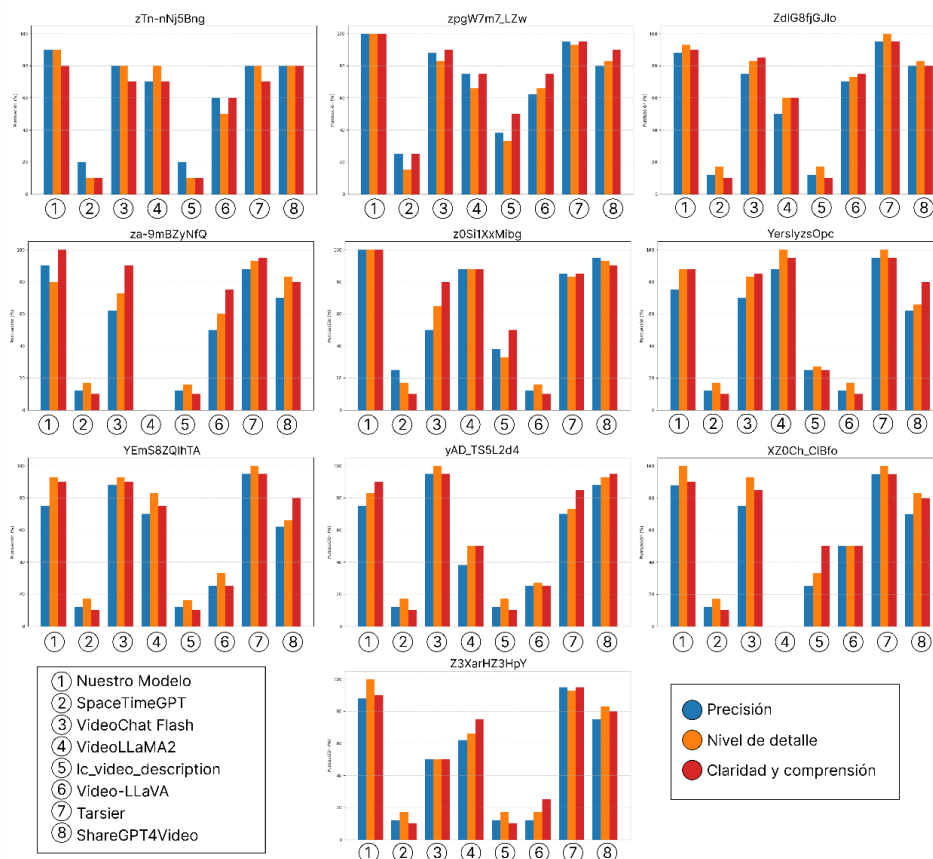


Fig. 2. Comparación de modelos de generación de descripciones de video en términos de precisión, nivel de detalle y claridad y comprensión.

6 Discusión

Los modelos evaluados presentan un desempeño variado en función de los tres criterios analizados. Se observa que nuestro modelo mantiene un desempeño alto en precisión, nivel de detalle y claridad y comprensión, destacándose como uno de los mejores modelos evaluados. En contraste, modelos como SpaceTimeGPT y Lc-video-description presentan un desempeño considerablemente bajo en casi todos los videos.

Los resultados obtenidos evidencian importantes diferencias en la calidad de las descripciones generadas por los distintos modelos evaluados. Se observaron patrones recurrentes en algunos modelos que impactaron directamente en los puntajes asignados a los criterios de precisión, nivel de detalle y claridad. En particular, se identificaron modelos cuya estructura de generación de descripciones seguía patrones repetitivos, limitando la variabilidad en la narrativa y afectando su capacidad para representar con precisión los eventos observados en los videos.

Caso contrario en modelos como Tarsier, ShareGPT4Video y nuestro modelo los cuales presentan una estructura narrativa más detallada y contextualizada, incorporando información relevante sobre el entorno y las acciones específicas de los personajes en escena. Estas descripciones lograron identificar correctamente los elementos principales del video introduciendo información complementaria que mejora la comprensión global de la escena.

Los resultados obtenidos en la caracterización prosódica reflejan una correlación entre las emociones predominantes en cada video y las características de la narración generada. Se analizaron parámetros clave como intensidad, tono, duración y ritmo, permitiendo una evaluación detallada de cómo la variabilidad emocional influye en la construcción de la audiodescripción.

Las emociones como alegría y éxtasis generan un tono alto, con variaciones amplias y un ritmo dinámico, mientras que tristeza se asocia con un tono monótono, pausas reflexivas y ritmo lento. Las descripciones relacionadas con confianza y aceptación presentan una intensidad estable y un tono medio, asegurando fluidez sin cambios bruscos. En contraste, emociones como anticipación y vigilancia muestran variabilidad en el tono y un ritmo ágil, reflejando mayor atención en la escena.

Estos resultados destacan la importancia de ajustar dinámicamente la prosodia en función de la emoción predominante, permitiendo generar audiodescripciones más naturales y alineadas con la intencionalidad emocional del contenido visual.

7 Limitaciones y trabajo futuro

Una de las principales limitaciones del modelo propuesto es que fue probado únicamente con videos sin audio. Esto implica que no se toma en cuenta la información sonora original del video, la cual podría complementar y enriquecer la generación de prosodia y la comprensión del contexto. Además, existe una limitación asociada al uso del conjunto de datos MSVD. Si bien este corpus es ampliamente utilizado, sus descripciones manuales tienden a ser muy breves, lo que condicionó la evaluación del modelo. A esto se suma que, tras una revisión de literatura, se identificaron 17 librerías orientadas a la descripción automática de video, pero al momento de su implementación, solo 7 pudieron ejecutarse correctamente debido a problemas relacionados con la instalación y compatibilidad de dependencias. Por tanto, solo se compararon los modelos de descripción automática que resultaron funcionales.

Se identificaron áreas de mejora, particularmente en la adaptabilidad prosódica en escenas de corta duración, donde la variabilidad emocional requiere ajustes más dinámicos en tono y ritmo. Así mismo, la implementación de estrategias avanzadas de fusión multimodal podría mejorar la precisión en escenarios complejos o con múltiples eventos simultáneos.

Otra línea pendiente de exploración es el ajuste dinámico del número de caracteres por segundo en las descripciones generadas. Si bien se estableció un límite proporcional a la duración del video, no se evaluó su impacto real en la comprensión y sincronización de las narraciones. Este aspecto es clave para garantizar que las audiodescripciones sean accesibles, comprensibles y estén adecuadamente alineadas con el ritmo visual del contenido.

Se recomienda continuar optimizando los mecanismos de integración multimodal y ajuste dinámico de la prosodia para consolidar el desempeño del modelo propuesto y mejorar la accesibilidad de las narraciones generadas.

8 Conclusiones

El modelo propuesto exhibió un desempeño notable en cuanto a precisión semántica, riqueza descriptiva y claridad narrativa, posicionándose por encima de otros enfoques en la evaluación realizada. Su capacidad para

generar narrativas contextualmente enriquecidas permitió ofrecer descripciones más completas, integrando información específica sobre los objetos, acciones y eventos presentes en cada video.

La incorporación de mecanismos avanzados de atención y prosodia adaptativa favoreció la producción de narraciones más naturales y alineadas emocionalmente con el contenido visual. A diferencia de otros modelos que mostraron precisión aceptable, el modelo propuesto logró una mayor riqueza prosódica y contextual, aportando una experiencia descriptiva más inmersiva y detallada.

La integración de características prosódicas en la audiodescripción representa un avance considerable hacia una accesibilidad más inclusiva, trascendiendo la mera precisión semántica para incorporar elementos emocionales que enriquecen la experiencia narrativa. Este enfoque dual permite no solo transmitir el contenido visual, sino también su carga emocional, aspecto fundamental para una comprensión integral del material audiovisual.

Los resultados obtenidos evidencian que la personalización prosódica basada en el contexto emocional de cada escena mejora sustancialmente la calidad percibida de las audiodescripciones. Esta adaptabilidad dinámica constituye un elemento diferenciador frente a enfoques tradicionales, permitiendo transiciones fluidas entre diferentes estados emocionales dentro de una misma narrativa.

Agradecimientos. Producto derivado del proyecto INV-ING-4048 financiado por la Universidad Militar Nueva Granada - Vigencia 2024. Los autores expresan su agradecimiento a la Universidad Militar Nueva Granada por el apoyo brindado. Asimismo, se reconoce la vinculación como profesor de tiempo completo del autor Christian Quintero a la Universidad Militar Nueva Granada, lo que contribuyó significativamente a la realización de esta investigación.

Declaración de intereses. Los autores declaran que no poseen conflicto de intereses que puedan ser relevantes para el contenido de este artículo.

Declaración sobre uso de IA: Este artículo ha sido apoyado con herramientas de inteligencia artificial, las cuales se emplearon para realizar revisiones gramaticales. Todas las decisiones finales sobre el contenido y la redacción fueron realizadas por los autores.

Referencias

1. D. Sharma, C. Dhiman, and D. Kumar, "Evolution of visual data captioning Methods, Datasets, and evaluation Metrics: A comprehensive survey," *Expert Syst Appl*, vol. 221, p. 119773, Jul. 2023, doi: 10.1016/j.eswa.2023.119773.
2. L. Zhou, Y. Zhou, J. J. Corso, R. Socher, and C. Xiong, "End-to-End Dense Video Captioning with Masked Transformer," Apr. 2018, [Online]. Available: <http://arxiv.org/abs/1804.00819>
3. Y. Li, T. Yao, Y. Pan, H. Chao, and T. Mei, "Jointly Localizing and Describing Events for Dense Video Captioning," Apr. 2018, [Online]. Available: <http://arxiv.org/abs/1804.08274>
4. M. Shen *et al.*, "Making Accessible Movies Easily: An Intelligent Tool for Authoring and Integrating Audio Descriptions to Movies," in *Proceedings of the 21st International Web for All Conference*, New York, NY, USA: ACM, May 2024, pp. 160–164. doi: 10.1145/3677846.3677855.
5. F. Huang, S. Zeng, J. Ke, S. Lei, and J. Wang, "A Video Description Model with Improved Attention Mechanism," *J Phys Conf Ser*, vol. 2384, no. 1, p. 012015, Dec. 2022, doi: 10.1088/1742-6596/2384/1/012015.
6. W. Deng, M. Zhu, M. Ma, and Y. Tian, "Analysis of Speech Prosody Characteristics of Teachers," *Lecture Notes in Education Psychology and Public Media*, vol. 34, no. 1, pp. 197–203, Jan. 2024, doi: 10.54254/2753-7048/34/20231974.
7. Y. Zhu and S. Jiang, "Attention-based Densely Connected LSTM for Video Captioning," in *Proceedings of the 27th ACM International Conference on Multimedia*, New York, NY, USA: ACM, Oct. 2019, pp. 802–810. doi: 10.1145/3343031.3350932.
8. J. Xie *et al.*, "AutoAD-Zero: A Training-Free Framework for Zero-Shot Audio Description," Jul. 2024, [Online]. Available: <http://arxiv.org/abs/2407.15850>
9. W. Zhao, X. Wu, and X. Zhang, "MemCap: Memorizing Style Knowledge for Image Captioning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, pp. 12984–12992, Apr. 2020, doi: 10.1609/aaai.v34i07.6998.
10. T. C. Smits, S. L'yi, A. P. Mar, and N. Gehlenborg, "AltGosling: Automatic Generation of Text Descriptions for Accessible Genomics Data Visualization," 2024, doi: 10.1093/bioinformatics/btae670/7900296.

11. D. Li, Y. Zhao, R. Cui, and L. Zhao, "Research on image text generation based on word2vec visual vocabulary attention," in *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, IEEE, Jan. 2021, pp. 344–348. doi: 10.1109/ACCTCS52002.2021.00075.
12. A. Radford *et al.*, "Learning Transferable Visual Models From Natural Language Supervision," 2021. [Online]. Available: <https://github.com/OpenAI/CLIP>.
13. A. Yang *et al.*, "Vid2Seq: Large-Scale Pretraining of a Visual Language Model for Dense Video Captioning," Feb. 2023, [Online]. Available: <http://arxiv.org/abs/2302.14115>
14. S. Islam, A. Dash, A. Seum, A. H. Raj, T. Hossain, and F. M. Shah, "Exploring Video Captioning Techniques: A Comprehensive Survey on Deep Learning Methods," *SN Comput Sci*, vol. 2, no. 2, p. 120, Apr. 2021, doi: 10.1007/s42979-021-00487-x.
15. G. Bertasius, H. Wang, and L. Torresani, "Is Space-Time Attention All You Need for Video Understanding?," Feb. 2021, [Online]. Available: <http://arxiv.org/abs/2102.05095>
16. X. Li *et al.*, "VideoChat-Flash: Hierarchical Compression for Long-Context Video Modeling," Dec. 2024, [Online]. Available: <http://arxiv.org/abs/2501.00574>
17. B. Lin *et al.*, "Video-LLaVA: Learning United Visual Representation by Alignment Before Projection," Nov. 2023, [Online]. Available: <http://arxiv.org/abs/2311.10122>
18. L. Yuan, J. Wang, H. Sun, Y. Zhang, and Y. Lin, "Tarsier2: Advancing Large Vision-Language Models from Detailed Video Description to Comprehensive Video Understanding," Jan. 2025, [Online]. Available: <http://arxiv.org/abs/2501.07888>
19. L. Chen *et al.*, "ShareGPT4Video: Improving Video Understanding and Generation with Better Captions," Jun. 2024, [Online]. Available: <http://arxiv.org/abs/2406.04325>
20. T. Nitta, T. Fukuzawa, and T. Tamaki, "Fine-grained length controllable video captioning with ordinal embeddings," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3506751.
21. Z. Cheng *et al.*, "VideoLLaMA 2: Advancing Spatial-Temporal Modeling and Audio Understanding in Video-LLMs," Jun. 2024, [Online]. Available: <http://arxiv.org/abs/2406.07476>

Síntesis de Voz Emocional Basada en Prosodia para Audiodescripción

Juan Sebastián Carballo Ñustes¹[0009-0008-5498-5024], Alexander Rozo-Torres¹[0000-0001-9343-2264],
Christian Quintero¹[0000-0001-5763-2373]

¹Universidad Militar Nueva Granada, Bogotá, Colombia
{est.juan.carballo, est.omar.rozo, chris-tian.quintero}@unimilitar.edu.co

Resumen. La accesibilidad en las audiodescripciones para personas con discapacidad visual requiere avances en la síntesis de voz que potencien la expresividad emocional y la adaptabilidad prosódica. Este trabajo presenta un enfoque que transforma descripciones textuales y parámetros prosódicos en audios expresivos, comparando tres estrategias de síntesis de voz: (1) texto a voz basado en análisis semántico, que ajusta la prosodia según el significado del contenido; (2) texto a voz controlado por prompts, que modifica entonación y ritmo mediante instrucciones explícitas; y (3) texto a voz basado en referencia, que replica la prosodia de audios preexistentes. El experimento se estructuró en tres fases: preparación de datos, generación de audiodescripciones y evaluación de características prosódicas. Se generaron 50 audios a partir de 10 descripciones utilizando cinco modelos de síntesis de voz. Las características prosódicas (intensidad, entonación, ritmo y duración de pausas) se analizaron con Parselmouth y OpenSMILE, comparando las señales acústicas con valores de referencia asociados a ocho emociones principales. Los resultados evidencian diferencias significativas en la capacidad de los modelos para expresar emociones a través de la prosodia. Este análisis resalta las fortalezas y limitaciones de cada modelo, contribuyendo al desarrollo de herramientas más inclusivas y expresivas para mejorar la accesibilidad de personas con discapacidad visual.

Palabras clave: Audiodescripción, Síntesis de Voz Emocional, Prosodia, Accesibilidad, Modelos de Texto a Voz.

1 Introducción

En la comunicación humana, la voz no solo transmite información, sino que también aporta intencionalidad, emoción y significado [4, 5]. La prosodia desempeña un papel clave en la percepción del habla, influyendo tanto en la inteligibilidad como en el impacto emocional del mensaje. En la narración integrada en la AD, resulta esencial para guiar al oyente, destacar elementos clave y transmitir intenciones emocionales durante el relato.

No obstante, los sistemas actuales de síntesis de voz, conocido en inglés como Text-to-Speech (TTS), presentan limitaciones en la generación de una prosodia natural y expresiva, lo que impacta negativamente en la inmersión del oyente y en la percepción de calidad del contenido [6]. En el caso de la AD, que es una herramienta esencial para garantizar la accesibilidad de las personas con discapacidad visual, la falta de variabilidad prosódica puede conducir a una experiencia auditiva monótona, limitando la transmisión de matices emocionales y, por consiguiente, afectando la comprensión y el disfrute del contenido por parte del usuario [7, 8].

Para abordar esta problemática, la literatura propone el desarrollo de modelos de síntesis de voz emocional (con sus siglas en inglés, ETTS), que buscan mejorar la expresividad del habla sintetizada [1, 9]. Algunos enfoques recurren a audios de referencia para replicar patrones prosódicos humanos, mientras que otros emplean prompts textuales y de referencia para modular la entonación del discurso [10][3]. Asimismo, técnicas basadas en aprendizaje profundo y representaciones latentes han permitido generar variaciones prosódicas más adaptativas y naturales [2].

Adicionalmente, el uso de prompts de lenguaje natural ha emergido como una técnica prometedora para controlar las emociones en el habla sintetizada [11]. Estudios recientes han demostrado que los prompts pueden transferir con precisión emociones desde el texto al habla, manteniendo la calidad de la voz y la identidad del hablante [10]. De igual modo, técnicas como la conversión de voz emocional (con sus siglas en inglés, EVC)

han permitido modificar el tono afectivo sin alterar la identidad del hablante, lo que resulta particularmente relevante en aplicaciones como la AD.

Desde una perspectiva teórica, modelos emocionales como el psicoevolutivo de Plutchik [12] y el circuplejo de Russell [13] han sido ampliamente utilizados para mapear la transmisión de emociones en el habla [14]. Estos modelos ofrecen marcos conceptuales sólidos para comprender cómo las emociones se estructuran y se expresan, proporcionando herramientas valiosas para el diseño de sistemas de síntesis de voz emocional. Sin embargo, su implementación en TTS sigue siendo un desafío debido a la complejidad de adaptar la prosodia a distintos estados emocionales y contextos narrativos [15].

En este sentido, el presente artículo propone la integración de métodos alternativos de modelado prosódico para mejorar la expresividad y adaptabilidad del habla sintetizada en narraciones y AD. Se exploran estrategias de generación de voz basadas en inferencia emocional y prosodia, evaluando su impacto en características prosódicas clave. Finalmente, se analizan las implicaciones de estos modelos en aplicaciones de accesibilidad y generación automatizada de contenidos narrativos, considerando tanto la literatura existente como las limitaciones actuales en validación emocional automatizada.

2 Trabajos relacionados

La síntesis de voz emocional, especialmente en aplicaciones de AD, ha sido abordada desde diversos enfoques metodológicos y técnicos. Un ejemplo es el trabajo de Cai et al. [16] que presentan un modelo que permite controlar la síntesis de voz emocional mediante una estrategia semi-supervisada. Inicialmente, entrenan un modelo de reconocimiento emocional en habla (SER) sobre un conjunto de datos sin etiquetas emocionales explícitas. Posteriormente, emplean este modelo para asignar automáticamente etiquetas emocionales a un corpus de TTS. Este enfoque posibilita la generación de habla con expresividad emocional específica y calidad acústica cercana al habla natural, reduciendo la necesidad de anotaciones manuales, lo cual resulta relevante en contextos con escasez de datos etiquetados.

Cornille et al. [17] desarrollan ConEx, un sistema de síntesis de voz basado en FastSpeech que proporciona control interactivo multinivel sobre la prosodia, combinando representaciones prosódicas globales y locales. Este método permite ajustes prosódicos específicos por fonema y ofrece al usuario la capacidad de interactuar iterativamente con el modelo, realizando modificaciones locales sin perder coherencia global. Esto resulta especialmente útil en AD, donde la adaptación prosódica a distintos contextos emocionales puede mejorar considerablemente la experiencia del usuario.

Choi y Hahn [18] proponen una EVC basada en redes secuencia a secuencia (seq2seq), que incorpora módulos de atención para mapear eficazmente la duración del habla emocional. Este modelo permite un control explícito de la intensidad emocional mediante la incorporación de un codificador emocional específico, proporcionando flexibilidad y precisión, crucial para captar matices emocionales de distinta intensidad en AD.

Kaneko y Kameoka [19] abordan la conversión de voz no paralela utilizando redes adversarias cíclicas (CycleGAN-VC), demostrando la capacidad de generar voz convertida con una variabilidad emocional sin la necesidad de datos paralelos. Su metodología resalta la importancia de técnicas adversarias en la síntesis de voz emocional para lograr resultados de alta calidad, evitando problemas como la sobre suavización de parámetros prosódicos, común en modelos estadísticos tradicionales.

Wu et al. [20] presentan un método de síntesis de voz emocional de extremo a extremo que adopta tokens de estilo global (con sus siglas en inglés, GST) para un entrenamiento semi-supervisado. Su metodología permite obtener una correspondencia directa entre los GST y las categorías emocionales con pocos datos etiquetados, facilitando una transferencia emocional precisa y efectiva. Este enfoque es especialmente útil en situaciones con recursos limitados, asegurando una expresividad emocional destacada con escasas anotaciones manuales.

Zhang et al. [21] introducen un enfoque basado en un autocodificador variacionales (VAE) para aprender representaciones latentes de estilos de habla en modelos TTS de extremo a extremo. Este método permite un control efectivo y transferencia de estilos mediante inferencia variacional, mejorando la generalización y la

flexibilidad en la síntesis de habla emocional. Los resultados demuestran una capacidad superior en transferencia no paralela de estilos, siendo relevante para aplicaciones que requieren alta adaptabilidad emocional.

En conjunto, estos trabajos demuestran la diversidad de enfoques técnicos para la síntesis de voz emocional, cada uno con ventajas particulares en calidad acústica, control prosódico, intensidad emocional y reducción de la dependencia de grandes volúmenes de datos etiquetados. Sin embargo, aún existe una brecha en la comparación sistemática de estas metodologías en el contexto específico de la audiodescripción en español. En este sentido, el presente trabajo contribuye al campo al evaluar y contrastar diferentes estrategias de síntesis de voz emocional aplicadas a la audiodescripción, analizando su capacidad para reproducir prosodia expresiva y su impacto en la accesibilidad para personas con discapacidad visual. Esta comparación proporciona evidencia empírica que orienta el desarrollo de sistemas más inclusivos y efectivos en aplicaciones de accesibilidad.

3 Esquema propuesto

El enfoque propuesto tiene como objetivo optimizar la expresividad emocional y la adaptabilidad prosódica en los sistemas TTS aplicados a la AD. La Fig. 1 **¡Error! No se encuentra el origen de la referencia.** ilustra el proceso de modelado prosódico, en el que las descripciones textuales y las características prosódicas se transforman en audios emocionales mediante diversas estrategias avanzadas de síntesis de voz. Este proceso permite ajustar parámetros como entonación, ritmo, intensidad y duración de pausas, mejorando tanto la naturalidad como la intención emocional del audio generado. Los componentes principales del esquema son los siguientes:

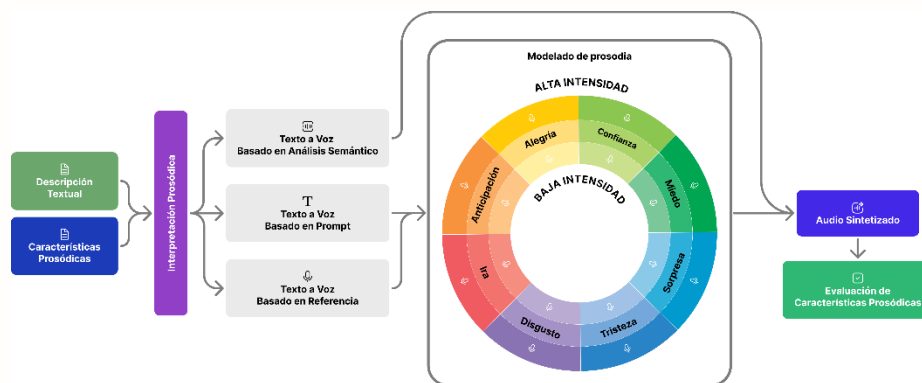


Fig. 1. Enfoque propuesto para el procesamiento de AD, integrando interpretación prosódica y síntesis de voz emocional.

Entrada: Descripción textual y características prosódicas. La entrada del sistema combina descripciones textuales y características prosódicas, que constituyen la base para la interpretación emocional. La descripción textual aporta el contenido semántico, mientras que las características prosódicas (entonación, ritmo, intensidad y duración de pausas) definen cómo debe ser expresado el mensaje. Este enfoque permite capturar tanto el significado del texto como la intención emocional.

Interpretación prosódica. En esta etapa, las entradas se transforman en parámetros prosódicos ajustables, lo que resulta fundamental para adaptar la prosodia a diferentes emociones y contextos narrativos. Este proceso aborda desafíos clave, como la naturalidad y la coherencia emocional, mediante el uso de técnicas avanzadas basadas en redes neuronales y modelos de aprendizaje profundo, que permiten mapear las emociones en términos prosódicos. Para lograr este objetivo, se implementan tres estrategias principales de síntesis de voz emocional:

- **Texto a voz basado en análisis semántico:** Ajusta la prosodia en función del significado del texto, identificando patrones de entonación a partir de palabras clave y estructuras gramaticales.

- **Texto a voz basado en prompt:** Permite configurar atributos prosódicos de manera explícita. Los usuarios pueden definir la emoción predominante, ajustar el nivel de énfasis y modificar el estilo narrativo según el propósito del mensaje.
- **Texto a voz basado en audio de referencia:** La prosodia se establece utilizando un conjunto de audios de entrada. Se extraen parámetros prosódicos que se direccionan hacia la intención emocional predominante.

Modelado de prosodia. El componente de modelado de prosodia tiene como objetivo establecer una correspondencia entre cada emoción definida en el modelo de Plutchik [22] y sus parámetros prosódicos característicos. Para ello, se recolectó un conjunto de audios de referencia que sirvieron como base para mapear perfiles prosódicos específicos a cada emoción.

Se elaboró un guion narrativo neutral, interpretado por tres locutores expertos en AD (dos voces masculinas y una voz femenina), cada uno realizando una lectura intencionada de las 16 emociones contempladas en la rueda emocional de Plutchik, organizadas en pares de intensidad alta y baja. Esto dio como resultado un conjunto de 48 grabaciones (3 locutores $\times$ 16 emociones).

Estas grabaciones fueron sometidas a un proceso de limpieza y depuración para eliminar ruidos no deseados y asegurar una calidad acústica uniforme a -12dB. Posteriormente, se analizaron automáticamente cuatro parámetros prosódicos fundamentales: frecuencia fundamental (F0), relacionada con la entonación; intensidad, que refleja la energía vocal; duración de pausas, como indicador de fluidez discursiva; y ritmo, entendido como la velocidad relativa del habla.

El análisis se realizó mediante un enfoque combinado utilizando Parselmouth (interfaz de PRAAT) y OpenSMILE, dos herramientas especializadas en el análisis acústico del habla que permiten una caracterización precisa y objetiva de la prosodia emocional. Los valores obtenidos se promediaron por emoción para construir un perfil prosódico representativo. Esta extracción no dependió de evaluación subjetiva directa; sin embargo, el guion fue previamente validado por el equipo de investigación para asegurar una interpretación emocional coherente durante la grabación. En la Fig. 2 se muestra el modelo prosódico derivado de las emociones interpretadas.

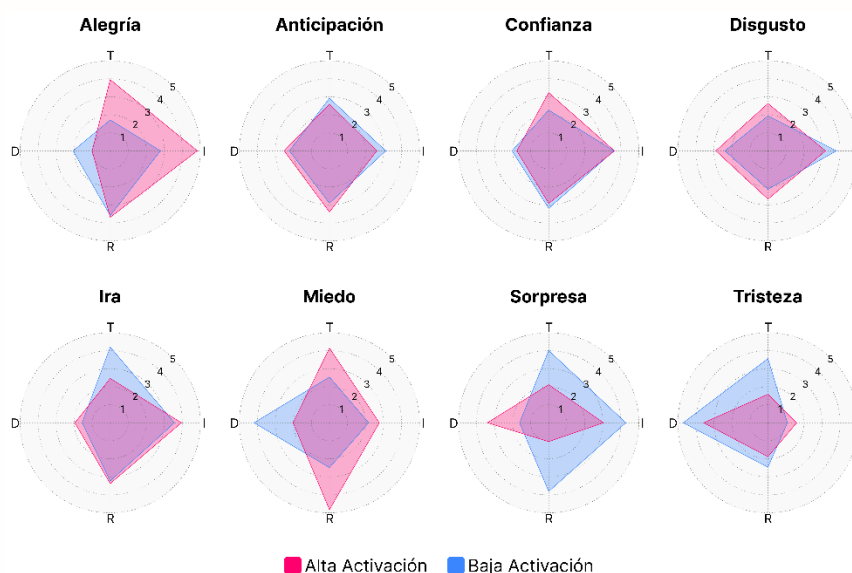


Fig. 2. Modelo prosódico basado en las 8 emociones básicas de Plutchik [22], derivadas de emociones secundarias opuestas, asociado a características prosódicas: intensidad (I), tono (T), duración (D) y ritmo (R).

Evaluación de características prosódicas. La evaluación prosódica constituye la fase final del esquema propuesto, en la que se analiza en qué medida el audio sintetizado refleja adecuadamente las emociones objetivo

a través de sus propiedades acústicas. Este proceso se basa en la comparación de las características prosódicas extraídas del audio generado (entonación, intensidad, duración de pausas y ritmo) con los valores de referencia previamente establecidos a partir de interpretaciones humanas expertas.

Esta evaluación no solo permite validar la capacidad de los modelos ETTS para reproducir variaciones emocionales en el habla, sino que también facilita la identificación de patrones específicos de desempeño y áreas de mejora en la síntesis prosódica, con implicaciones directas para aplicaciones de accesibilidad como la audiodescripción automatizada.

4 Metodología Experimental

El objetivo de la experimentación fue evaluar la calidad prosódica y la estabilidad de distintos modelos de síntesis de voz en la generación de AD. Para ello, se implementó un procedimiento estructurado en tres fases: preparación de datos, generación de AD y evaluación de características prosódicas, siguiendo el enfoque propuesto.

4.1 Preparación de datos

Se utilizó el conjunto de datos Microsoft Research Video Description Corpus (MSVD) [23], compuesto por videos cortos sin audio original, subtítulos ni texto superpuesto. Los videos tienen una duración promedio de entre 10 y 25 segundos y suelen centrarse en una única actividad, lo que facilita su segmentación y análisis. Para este estudio, se seleccionaron 10 videos previamente procesados en trabajos relacionados con la generación de narrativas semánticas y características prosódicas aplicadas a la audiodescripción en español. Esta selección permitió aprovechar insumos validados en términos semánticos y prosódicos, asegurando coherencia metodológica para la comparación entre los modelos evaluados.

La selección consideró una diversidad representativa de escenas, estilos narrativos y contextos visuales, con el fin de evaluar exhaustivamente el desempeño de los modelos de síntesis de voz en diferentes escenarios discursivos. El proceso experimental se organizó en varias fases, desde la reutilización de descripciones textuales existentes hasta la síntesis de voz emocional y el posterior análisis acústico.

Para cada video, se empleó un sistema automático de descripción visual (captioning) que proporcionó tanto la descripción narrativa del contenido visual, utilizada como entrada para generar la AD, como la extracción de características prosódicas clave (entonación, ritmo, pausas e intensidad). Estas características se obtuvieron con el propósito de mejorar la precisión emocional de los audios generados por los modelos de síntesis de voz.

4.2 Generación de AD

En la etapa de síntesis de voz, se implementó el esquema propuesto, aplicando las tres estrategias de generación descritas: análisis semántico, prompts y referencia. Estas estrategias se llevaron a cabo mediante cinco modelos de ETTS seleccionados por su diversidad de enfoques, compatibilidad con personalización prosódica y disponibilidad. La Tabla 1 sintetiza las principales características de cada herramienta.

Tabla 1. Herramientas de síntesis de voz y sus características.

Categoría	Herramienta	Descripción
Modelo autoregresivo general	Bark	Desarrollado por Suno, este modelo busca replicar la prosodia natural del habla. Genera voz en múltiples idiomas e incluye elementos como risas y pausas, lo que mejora la naturalidad.

Texto a voz basado en prompt	Parler-TTS	Genera voz de alta calidad con control mediante instrucciones textuales que permiten modificar género, velocidad, tono, ruido de fondo y reverberación.
Texto a voz basado en prompt	GPT-4o (preview)	Modelo flexible de OpenAI que permite un control preciso sobre prosodia mediante prompts. Permite ajustar acento, emoción, velocidad, tono y género. Soporta más de 50 idiomas, cubriendo el 97% de los hablantes del mundo.
Texto a voz basado en audio de referencia	Coqui	Plataforma de código abierto que genera voz natural a partir de ejemplos de referencia. Compatible con múltiples idiomas y ofrece herramientas de entrenamiento y personalización.
Texto a voz basado en audio de referencia	Spanish-F5	Modelo basado en F5-TTS, entrenado con 218 horas de audio en español de varios países. Genera voz emotiva y natural con soporte para acentos regionales.

¡Cada modelo fue empleado para sintetizar AD a partir de las descripciones textuales generadas para los 10 videos del corpus experimental. En todos los casos, el contenido semántico de las descripciones se mantuvo constante, permitiendo que las diferencias observadas en la salida auditiva respondieran exclusivamente a las variaciones prosódicas introducidas conforme al enfoque propuesto.

La configuración de cada herramienta respetó los lineamientos propios de su categoría funcional: en los modelos basados en análisis semántico (como Bark), la prosodia se derivó implícitamente de la estructura lingüística del texto; en los modelos controlados mediante prompts (como GPT-4o y Parler-TTS), las emociones objetivo se especificaron de manera explícita mediante instrucciones textuales; y en los modelos por referencia (Coqui y Spanish-F5), la prosodia fue replicada a partir de grabaciones humanas previamente clasificadas según sus características acústicas. A continuación, se muestra el tipo de entrada utilizada para cada modelo:

Bark: Descripción textual + Características prosódica.

Parler-TTS: Descripción textual + Características prosódicas.

GPT-4o: Se utilizó un esquema jerárquico de instrucciones en formato de roles de conversación, con un rol de sistema que define el comportamiento global del sintetizador y un rol de usuario con el contenido a leer. El prompt empleado fue el siguiente:

```
{ "role": "system", "content": "
You are an advanced text-to-speech assistant, specialized in delivering natural and emotionally expressive speech.
Your task is to convert the given text into spoken audio in Colombian Spanish, ensuring an authentic and engaging reading experience.
- Emotional expression: Read the text with a {emotion} tone, capturing and conveying the sentiment effectively and expressly.
- Prosodic adaptation: Convey the emotion of {emotion} with the following characteristics: {prosodic_characteristics}.
- Clarity and authenticity: Maintain clear pronunciation while preserving the nuances of Colombian Spanish.
Important: Do not alter the content of the text in any way." },
{ "role": "user", "content": "text" }
```

Coqui TTS: Interpretación pregrabada + Descripción textual

Spanish-F5: Interpretación pregrabada + Descripción textual

Para los modelos Bark, GPT-4o y Parler-TTS, se generó una audiodescripción por video, totalizando 30 archivos de audio (3 modelos $\times$ 10 descripciones). Los modelos por referencia fueron utilizados con tres locutores expertos en AD (dos masculinos y una femenina), lo que permitió generar tres versiones prosódicas diferenciadas por video y por modelo. Esto resultó en 60 audios adicionales (2 modelos $\times$ 3 locutores $\times$ 10 descripciones). En conjunto, el corpus final constó de 90 ADs sintetizadas, cuidadosamente balanceadas para permitir comparaciones tanto inter-modelo como inter-locutor.

Este corpus se utilizó posteriormente como base para el análisis comparativo de características prosódicas, evaluando el grado de alineación entre los parámetros generados y los perfiles emocionales definidos. La sistematización del proceso de síntesis, junto con el control sobre las variables semánticas y prosódicas, permitió establecer condiciones homogéneas para una evaluación rigurosa del impacto emocional en la narración automatizada.

4.3 Evaluación de las características prosódicas

El análisis de las ADs sintetizadas se realizó utilizando la herramienta OpenSMILE, reconocida por su robustez en la extracción automatizada de descriptores acústicos a partir de señales de voz. Esta librería resultó especialmente adecuada para el presente estudio por su capacidad de procesar grandes volúmenes de datos de manera eficiente, garantizando así una caracterización sistemática y reproducible de las propiedades prosódicas de cada muestra.

La evaluación se centró en cuatro parámetros fundamentales de la prosodia, seleccionados por su relevancia en la percepción emocional del habla:

- **Frecuencia fundamental (F0):** Indicador de la variación en entonación.
- **Duración de pausas:** Relacionada con el ritmo y la fluidez del habla.
- **Intensidad:** Reflejo de la energía en la voz y su impacto perceptivo.
- **Ritmo:** Velocidad del habla, considerando la duración de sílabas y pausas.

En el caso de los modelos de síntesis prosódica, que generaron más versiones de AD por descripción, se extrajeron los valores prosódicos de cada muestra y se calcularon los promedios correspondientes. Este procedimiento permitió consolidar una única interpretación prosódica por modelo y video, facilitando la comparación directa con los demás sistemas evaluados y construyendo un corpus homogéneo y equilibrado de características acústicas, sobre el cual se evaluó la prosodia sintetizada.

5 Resultados

Para analizar la expresividad emocional de las ADs generadas, se realizó una evaluación detallada de las salidas de los cinco modelos de síntesis de voz implementados. La valoración se centró en la capacidad de cada sistema para reproducir los patrones prosódicos asociados a las emociones objetivo, considerando los cuatro parámetros fundamentales: intensidad (I), tono (T), ritmo (R) y duración de pausas (D).

En total, se evaluaron 50 ADs, cada una vinculada a una emoción básica del modelo de Plutchik, contemplando niveles de activación alta y baja. La **Fig. 1** presenta una visualización comparativa del grado de alineación emocional entre los valores generados por los modelos y los perfiles prosódicos de referencia. Por su parte, la **Tabla 1** resume los promedios obtenidos para cada parámetro prosódico, así como los errores cuadráticos medios (RMSE) respecto a los valores esperados.

Tabla 2. Promedios y errores RMSE por modelo de síntesis de voz.

Modelo	Promedios				RMSE			
	I	T	D	R	I	T	D	R
BarkTTS	3.7	3.5	1.8	3.0	1.30	1.40	0.99	1.10
Coqui	3.6	2.7	2.1	3.8	0.74	0.70	0.71	0.69
GPT-4o	2.4	1.7	2.1	3.4	1.50	1.40	0.66	0.59
ParlerTTS	2.8	2.3	2.9	2.1	1.50	1.10	1.60	1.50
SpanishF5	2.3	3.1	1.8	4.0	1.50	0.79	0.44	0.76

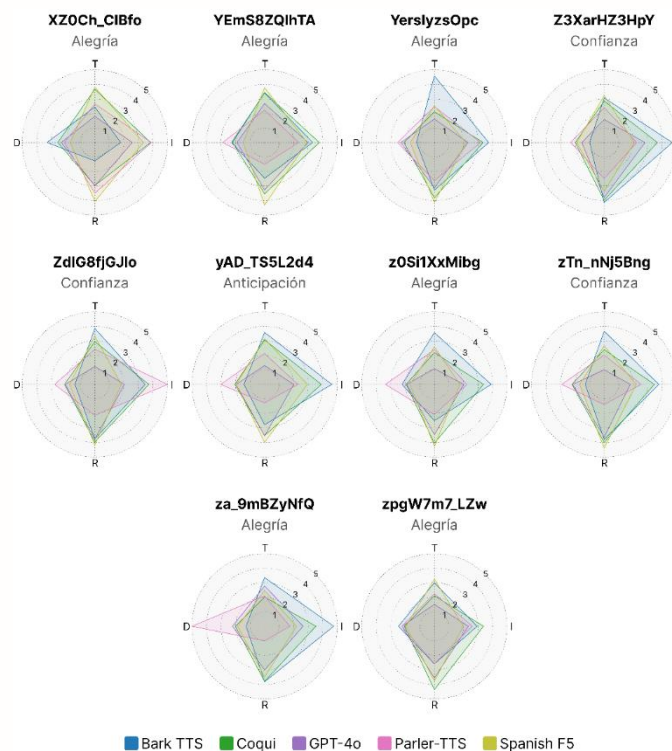


Fig. 3. Comparación de los parámetros prosódicos generados por los modelos de síntesis de voz frente a los valores esperados según el modelo emocional de Plutchik. La visualización representa el ajuste medio en intensidad (I), tono (T), ritmo (R) y duración de pausas (D), evaluado mediante OpenSMILE en las 50 audiodescripciones.

6 Discusión

El análisis de la Fig.3 revela diferencias notables en el desempeño prosódico de los modelos evaluados. Coqui destaca por presentar los valores promedio más equilibrados y los menores errores RMSE en la mayoría de los parámetros, especialmente en intensidad, tono y ritmo. Spanish-F5, por su parte, muestra un rendimiento sobresaliente en el ajuste del ritmo (R) y la duración de pausas (D), con los menores errores en estos parámetros, lo que sugiere una mayor capacidad para replicar la fluidez y velocidad del habla emocional.

En contraste, BarkTTS y GPT-4o tienden a presentar mayores desviaciones respecto a los valores de referencia, particularmente en intensidad y tono, lo que indica una menor precisión en la reproducción de

matices emocionales. ParlerTTS, aunque logra un buen ajuste en la duración de pausas, muestra una mayor variabilidad en el ritmo y la intensidad, lo que puede afectar la coherencia prosódica.

El análisis de casos específicos permite ilustrar mejor estas diferencias. Por ejemplo, en la AD correspondiente a *zTn_nNj5Bng* (Alegria), GPT-4o Audio logró mayor expresividad al modular intensidad y entonación con dinamismo, mientras que Spanish-F5 y Coqui ofrecieron interpretaciones más neutras y BarkTTS careció de variabilidad tonal. En *zpgW7m7_LZw* (Anticipación), los modelos basados en prompts capturaron la modulación ascendente y el ritmo ágil esperado, mientras que los de referencia presentaron menor dinamismo y BarkTTS mostró baja intensidad.

Para *ZdlG8ffGJlo* (Éxtasis), los modelos controlados por prompts lograron una entonación amplia y energía elevada, mientras que los de referencia fueron más contenidos y BarkTTS presentó fluctuaciones erráticas. En *za_9mBZyNfQ* (Serenidad), Spanish-F5 y BarkTTS ofrecieron estabilidad en ritmo y pausas, mientras que GPT-4o Audio resultó más expresivo de lo ideal y los demás modelos presentaron variabilidad tonal mayor a la esperada. Finalmente, en *YerslyzsOpc* (Éxtasis), ParlerTTS alcanzó los valores más cercanos a la emoción esperada, mientras que los modelos de referencia mostraron menor intensidad y BarkTTS resultó más neutro.

La Tabla 2 permite observar con precisión el error cuadrático medio (RMSE) por característica prosódica, facilitando la evaluación de la cercanía de cada modelo a los perfiles ideales de intensidad, entonación, duración y ritmo. Coqui se posiciona como el modelo con mayor precisión prosódica global, registrando los RMSE más bajos en casi todas las dimensiones analizadas (intensidad = 0.74, tono = 0.70, duración = 0.71 y ritmo = 0.69). Esta consistencia indica que Coqui es particularmente eficiente generando patrones prosódicos coherentes, acercándose significativamente al perfil emocional deseado, lo que lo convierte en una opción preferente para aplicaciones que requieren prosodia expresiva y estable en contextos emocionales diversos.

Spanish-F5 también presenta un desempeño destacado, con valores promedio alineados con los ideales y un RMSE mínimo en duración (0.44). Esta precisión temporal indica que reproduce con alta fidelidad los patrones rítmicos y pausas asociados a la expresión emocional, contribuyendo a transmitir matices afectivos en el habla sintetizada. Además, su ritmo promedio elevado podría favorecer una percepción más dinámica o enérgica del discurso, especialmente útil en emociones de alta activación.

Por otro lado, ParlerTTS presenta mayor desviación en el RMSE en múltiples dimensiones, lo que puede comprometer la coherencia prosódica y generar una expresividad emocional menos fiable y potencialmente confusa para el usuario final.

Finalmente, BarkTTS, basado en un enfoque semántico, mostró el desempeño más débil de los cinco modelos evaluados, con los valores más elevados de RMSE en intensidad (1.30), tono (1.40), duración (0.99) y ritmo (1.10). Este modelo exhibe una prosodia menos precisa, con valores dispersos y un mayor alejamiento de los perfiles emocionales esperados, posiblemente debido a su menor especialización en control emocional explícito.

En síntesis, los resultados evidencian que la estrategia de referencia humana (Coqui y Spanish-F5) sigue siendo la más robusta para la reproducción fiel de prosodia emocional en audiodescripción, mientras que los modelos basados en prompts ofrecen ventajas en emociones de alta energía, y los modelos semánticos requieren mejoras para alcanzar una mayor expresividad y precisión emocional.

7 Limitaciones y trabajo a futuro

Este trabajo presenta una serie de limitaciones que deben considerarse para mejorar la interpretación de los resultados y la planificación de investigaciones futuras.

En primer lugar, el tamaño y la representatividad de la muestra constituyen una restricción importante. El análisis se basó en un conjunto reducido de 10 videos del corpus MSVD, lo que, si bien permitió un control experimental riguroso y un análisis detallado, limita la generalización de los hallazgos a otros contextos, géneros narrativos o públicos. La selección específica de los videos puede no capturar toda la variabilidad prosódica y emocional presente en situaciones reales de audiodescripción.

Otra limitación relevante es la ausencia de validación perceptual con usuarios humanos, especialmente personas con discapacidad visual, quienes constituyen el público objetivo de las ADs generadas. La evaluación de las emociones sintetizadas se realizó exclusivamente mediante métricas acústicas extraídas con herramientas automáticas, sin aplicar validaciones subjetivas que permitan contrastar la percepción humana de la emoción expresada. Esta omisión podría afectar la precisión en la asignación emocional o la naturalidad percibida en los audios.

Desde el punto de vista técnico, los modelos utilizados presentan restricciones intrínsecas. No se consideraron explícitamente contextos prosódicos extendidos o discursos prolongados, en los cuales las emociones pueden experimentar transiciones o fluctuaciones que afectan la coherencia narrativa. Ahora bien, la dependencia de herramientas automáticas para la extracción y análisis de características prosódicas puede introducir sesgos o limitaciones en la interpretación de los resultados, especialmente en lo que respecta a la correspondencia entre los parámetros acústicos y la percepción emocional real.

Estas limitaciones abren diversas oportunidades para investigaciones futuras. Una línea prioritaria será ampliar el corpus de evaluación, incorporando videos con mayor diversidad temática y narrativa, así como muestras más extensas, para mejorar la generalización y robustez del esquema propuesto. Además, se propone realizar validaciones perceptuales con usuarios finales, especialmente personas con discapacidad visual, a fin de contrastar las métricas acústicas con la percepción subjetiva de expresividad y naturalidad.

Otra dirección clave es el desarrollo de modelos dinámicos de prosodia, capaces de adaptar la emoción de forma continua en narraciones más extensas, simulando cambios afectivos a lo largo del discurso. También será relevante explorar técnicas avanzadas para mitigar el sobreajuste observado en modelos como Spanish-F5, e implementar mecanismos que permitan preservar la fidelidad del texto original en modelos como GPT-4o.

Se plantea, además, la posibilidad de desarrollar sistemas híbridos que combinen la estabilidad y claridad prosódica de modelos como GPT-4o con la expresividad y naturalidad emocional de Spanish-F5 y Coqui, abordando así las limitaciones específicas de cada enfoque. Finalmente, la integración de información multimodal —que relacione voz, escena y narrativa— podría enriquecer la coherencia emocional y la calidad expresiva de las ADs prosódicas generadas, fortaleciendo la aplicabilidad del enfoque propuesto en escenarios reales de accesibilidad.

8 Conclusiones

El análisis comparativo de los modelos de síntesis de voz evaluados evidencia diferencias significativas en estabilidad, expresividad y fidelidad a las instrucciones prosódicas. Los resultados muestran que algunos modelos presentan limitaciones en la modulación prosódica y estabilidad del audio, mientras que otros logran un equilibrio entre expresividad y consistencia, con implicaciones clave para aplicaciones de audiodescripción (AD) en el campo de la accesibilidad.

Bark y Parler-TTS presentaron deficiencias notables, generando prosodias planas con entonación monótona y errores recurrentes en dicción, silencios abruptos y ruidos inconsistentes, lo que afecta la claridad y coherencia del mensaje. GPT-4o demostró mayor estabilidad y precisión en la modulación prosódica, aunque con expresividad limitada, lo que lo hace adecuado para aplicaciones donde la claridad es prioritaria. Sin embargo, se observó que GPT-4o puede modificar ligeramente el contexto del texto proporcionado, ajustando el contenido leído para coincidir con la emoción requerida, lo que en ocasiones puede alterar el sentido original del mensaje.

Coqui destacó por su emocionalidad reconocible y ausencia de errores significativos, aunque presentó pequeñas congelaciones en audios largos, lo que podría afectar la fluidez en narraciones extensas. Por su parte, Spanish-F5 generó las voces más naturales y expresivas, con una modulación prosódica efectiva para transmitir emociones. No obstante, presentó problemas de estabilidad en la dicción y generación de contenido no intencionado, así como la filtración del audio de referencia en el audio generado, lo que sugiere un posible sobreajuste del modelo.

En síntesis, los resultados subrayan la necesidad de optimizar tanto la estabilidad como la expresividad de los modelos de síntesis de voz para aplicaciones de AD, priorizando la coherencia y naturalidad del audio generado. Mejorar estos aspectos será fundamental para garantizar una experiencia accesible, comprensible y emocionalmente enriquecida para las personas con discapacidad visual, contribuyendo así al desarrollo de tecnologías más inclusivas y efectivas en el ámbito de la accesibilidad.

Agradecimientos. Producto derivado del proyecto INV-ING-4048 financiado por la Universidad Militar Nueva Granada - Vigencia 2024. Los autores expresan su agradecimiento a la Universidad Militar Nueva Granada por el apoyo brindado. Asimismo, se reconoce la vinculación como profesor de tiempo completo del autor Christian Quintero a la Universidad Militar Nueva Granada, lo que contribuyó significativamente a la realización de esta investigación.

Declaración de intereses. Los autores declaran que no poseen conflicto de intereses que puedan ser relevantes para el contenido de este artículo.

Declaración sobre uso de IA. Este artículo ha sido apoyado con herramientas de inteligencia artificial, las cuales se emplearon para realizar revisiones gramaticales. Todas las decisiones finales sobre el contenido y la redacción fueron realizadas por los autores.

Referencias

1. B. M. Ben-David, D. R. Chebat, and M. Icht, "'Love looks not with the eyes': supranormal processing of emotional speech in individuals with late-blindness versus preserved processing in individuals with congenital-blindness," *Cogn Emot*, 2024, doi: 10.1080/02699931.2024.2357656.
2. B. Iavarone et al., "The linguistic structure of an emotional text influences the sympathetic activity and the speech prosody of the reader," *Biomed Signal Process Control*, vol. 89, no. November 2023, p. 105776, 2024, doi: 10.1016/j.bspc.2023.105776.
3. Z. Gu and K. He, "Affective Prompt-Tuning-Based Language Model for Semantic-Based Emotional Text Generation," *Int J Semant Web Inf Syst*, vol. 20, no. 1, 2024, doi: 10.4018/IJSWIS.339187.
4. Y.-L. Liu, Y.-X. Zhang, Y. Wang, and Y. Yang, "Evidence for early encoding of speech in blind people," *Brain Lang*, vol. 259, p. 105504, 2024, doi: <https://doi.org/10.1016/j.bandl.2024.105504>.
5. C. Nussbaum, M. Poehlmann, H. Kreysa, and S. R. Schweinberger, "Perceived naturalness of emotional voice morphs," *Cogn Emot*, vol. 37, no. 4, pp. 731–747, May 2023, doi: 10.1080/02699931.2023.2200920.
6. T. M. Taha, Z. Ben Messaoud, and M. Frikha, "Convolutional Neural Network Architectures for Gender, Emotional Detection from Speech and Speaker Diarization," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 3, pp. 88–103, 2024, doi: 10.3991/ijim.v18i03.43013.
7. K. Cusworth, G. Paulik, N. Thomas, D. Preece, G. Campitelli, and D. C. Mathersul, "Emotion processes in voice-hearers: Understanding differences in emotional reactivity, emotion regulation and alexithymia," *Psychology and Psychotherapy: Theory, Research and Practice*, vol. 97, no. 4, pp. 706 – 721, 2024, doi: 10.1111/papt.12554.
8. X. Qi, C. Liu, L. Li, J. Hou, H. Xin, and X. Yu, "EmotionGesture: Audio-Driven Diverse Emotional Co-Speech 3D Gesture Generation," *IEEE Trans Multimedia*, vol. 14, no. 8, pp. 1–10, 2024, doi: 10.1109/TMM.2024.3407692.
9. J. Sarzedas, C. F. Lima, M. S. Roberto, S. K. Scott, A. P. Pinheiro, and T. Conde, "Blindness influences emotional authenticity perception in voices: Behavioral and ERP evidence," *Cortex*, vol. 172, no. xxxx, pp. 254–270, 2024, doi: 10.1016/j.cortex.2023.11.005.
10. T. Bott, F. Lux, and N. T. Vu, "Controlling Emotion in Text-to-Speech with Natural Language Prompts," Jun. 2024, [Online]. Available: <http://arxiv.org/abs/2406.06406>.
11. X. Zheng, Y. Du, and X. Qin, "CoMaSa: Context Multi-aware Self-attention for emotional response generation," *Neurocomputing*, vol. 611, 2025, doi: 10.1016/j.neucom.2024.128692.
12. R. PLUTCHIK, *Measuring Emotions and Their Derivatives*, vol. 4. Academic Press, Inc., 1989. doi: 10.1016/b978-0-12-558704-4.50007-4.
13. J. A. Russell, J. Clement, D. Jiwani, K. Ridgeway, and M. Schroeder, "A Circumplex Model of Affect," Ryman, 1980.
14. Z. Zhou, Y. Shen, X. Chen, and D. Wang, "GERP: A Personality-Based Emotional Response Generation Model," *Applied Sciences (Switzerland)*, vol. 13, no. 8, pp. 1–15, 2023, doi: 10.3390/app13085109.
15. H. S. Oh, S. H. Lee, and S. W. Lee, "DiffProsody: Diffusion-Based Latent Prosody Generation for Expressive Speech Synthesis With Prosody Conditional Adversarial Training," *IEEE/ACM Trans Audio Speech Lang Process*, vol. 32, pp. 2654–2666, 2024, doi: 10.1109/TASLP.2024.3395994.
16. X. Cai, D. Dai, Z. Wu, X. Li, J. Li, and H. Meng, "Emotion controllable speech synthesis using emotion-unlabeled dataset with the assistance of cross-domain speech emotion recognition," Oct. 2020, doi: <https://doi.org/10.48550/arXiv.2010.13350>.

17. T. Cornille, F. Wang, and J. Bekker, "Interactive Multi-Level Prosody Control for Expressive Speech Synthesis," in ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, May 2022, pp. 8312–8316. doi: 10.1109/ICASSP43922.2022.9746654.
18. H. Choi and M. Hahn, "Sequence-to-Sequence Emotional Voice Conversion With Strength Control," IEEE Access, vol. 9, pp. 42674–42687, 2021, doi: 10.1109/ACCESS.2021.3065460.
19. T. Kaneko and H. Kameoka, "CycleGAN-VC: Non-parallel Voice Conversion Using Cycle-Consistent Adversarial Networks," in 2018 26th European Signal Processing Conference (EUSIPCO), IEEE, Sep. 2018, pp. 2100–2104. doi: 10.23919/EUSIPCO.2018.8553236.
20. P. Wu, Z. Ling, L. Liu, Y. Jiang, H. Wu, and L. Dai, "End-to-End Emotional Speech Synthesis Using Style Tokens and Semi-Supervised Training," Jun. 2019, doi: <https://doi.org/10.48550/arXiv.1906.10859>.
21. Y.-J. Zhang, S. Pan, L. He, and Z.-H. Ling, "Learning latent representations for style control and transfer in end-to-end speech synthesis," Dec. 2018, [Online]. Available: <http://arxiv.org/abs/1812.04342>.
22. R. Plutchik, "The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," Am Sci, vol. 89, no. 4, pp. 344–350, 2001, [Online]. Available: <http://www.jstor.org/stable/27857503>.
23. S. Islam, A. Dash, A. Seum, A. H. Raj, T. Hossain, and F. M. Shah, "Exploring Video Captioning Techniques: A Comprehensive Survey on Deep Learning Methods," SN Comput Sci, vol. 2, no. 2, p. 120, Apr. 2021, doi: 10.1007/s42979-021-00487-x.



CONSEJO ZACATECANO DE
CIENCIA, TECNOLOGÍA
E INNOVACIÓN
ESTADO DE ZACATECAS



HCI
COLLAB