

INTELIGENCIA ARTIFICIAL APLICADA AL RETO DEMOGRÁFICO

*Ana Belén Gil González, Laura Grande Pérez,
Javier Parra Domínguez (coords.)*



Ediciones Universidad
Salamanca

Inteligencia artificial aplicada al reto demográfico



Colabora



Coordinadores

Ana Belén Gil González
Laura Grande Pérez
Javier Parra Domínguez

Comité organizador

Ana Belén Gil González
Javier Parra Domínguez
Jesús Ángel Román Gallego
Juan Manuel Núñez Velasco
Laura Grande Pérez
Joysiane Monroy
Zoe Doncel Cruz

Comité científico

Adrian Iftene
Ana Belén Gil González
Covadonga Rodrigo San Juan
Javier Parra Domínguez
Jesús Ángel Román Gallego
Juan Manuel Núñez Velasco
Rafael Capilla Sevilla
Laura Grande Pérez
Sara Rodríguez González
Joysiane Monroy

ANA BELÉN GIL GONZÁLEZ, LAURA GRANDE PÉREZ,
JAVIER PARRA DOMÍNGUEZ (COORDS.)

Inteligencia artificial aplicada al reto demográfico



Ediciones Universidad
Salamanca

AQUILAFUENTE, 394

© Ediciones Universidad de Salamanca y los autores

1ª edición: junio, 2026

ISBN: 978-84-1091-189-5 (pdf)
978-84-1091-190-1 (Impreso POD)

DOI: <https://doi.org/10.14201/0AQ0394>

Ediciones Universidad de Salamanca
Plaza San Benito, 2
E-37002 Salamanca (España)
<http://www.eusal.es>
eus@usal.es

Composición y realización:
Gráficas Lope
www.graficaslope.com

Esta publicación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia.

Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

Hecho en UE-Made in EU

Obra evaluada por el sistema de doble ciego determinado
por el Comité Científico del Congreso.



Usted es libre de: Compartir – copiar y redistribuir el material en cualquier medio o formato
Ediciones Universidad de Salamanca no revocará mientras cumpla con los términos:

- Reconocimiento – Debe reconocer adecuadamente la autoría, proporcionar un enlace a la licencia e indicar si se han realizado cambios. Puede hacerlo de cualquier manera razonable, pero no de una manera que sugiera que tiene el apoyo del licenciador o lo recibe por el uso que hace.
- NoComercial – No puede utilizar el material para una finalidad comercial.
- SinObraDerivada – Si remezcla, transforma o crea a partir del material, no puede difundir el material modificado.

Ediciones Universidad de Salamanca es miembro de la UNE
Unión de Editoriales Universitarias Españolas www.une.es



Catalogación de editor en ONIX accesible en <https://www.dilve.es>

ÍNDICE

Introducción.....	11
Dra. Ana BELÉN GIL	
1. COMUNICACIONES ORALES.....	13
1.1. CONTRIBUCIONES CIENTÍFICAS.....	15
Aplicación de inteligencia artificial para predecir el impacto del cambio climático en la fenología de la vid en Borgoña (Francia): implicaciones para la sostenibilidad vitivinícola..	17
Sofania M. ROJAS-LANDACAY, Juan M. NÚÑEZ V., José-Abel FLORES y Fernando DE LA PRIETA	
Transformaciones en la educación superior: redefinición del papel del profesorado ante la aparición de la inteligencia artificial generativa.....	23
Claudia JIMENO POSTIGO	
Construcción y validación de una ontología para el desarrollo de soluciones de inteligencia artificial orientadas al reto demográfico.....	29
Jorge DOMÍNGUEZ	
CytoLens: marco preliminar para el apoyo al cribado y la priorización de hallazgos en citología cervical digital mediante inteligencia artificial.....	35
Sandra MARCOS RECIO, Lautaro ROSSI LABIANCA, Ana-Belén GIL-GONZÁLEZ y Andrés CARDONA-MENDOZA	
DataMarket: aplicación de analítica inteligente para la optimización y la fidelización en minimarkets de barrio.....	41
Juan Manuel LÓPEZ, Juan Felipe ZAMBRANO, Juan CAMILO BURBANO, Andrés ENRÍQUEZ y Juan M. NÚÑEZ V.	
Modelización inteligente para la gestión sostenible de biomasa forestal mediante IA geoespacial.....	49
Sarai COBAS MERINO, Darío LÓPEZ RAMOS, Jairo CALDERÓN MICAÑA y Ana-Belén GIL-GONZÁLEZ	
La inteligencia artificial (IA) en la asistencia a personas mayores en entornos rurales: revisión del estado del arte.....	55
Laura GRANDE-PÉREZ, Johanna MORENO-CASTELLANOS y Ana Belén GIL-GONZÁLEZ	
Fundamentos y diseño de un servicio basado en inteligencia artificial para el bienestar financiero y farmacológico.....	61
José ORTEGA-MOHEDANO y Jaime SUÁREZ PÉREZ	
Una herramienta unificada para estudiar los cultivos forestales de España.....	67
Jaime GABRIEL VEGAS, Mairder Araceli URBON JIMÉNEZ , Ana DE LUIS REBOREDO, Belén PÉREZ LANCHO y Ana-Belén GIL-GONZÁLEZ	
Optimización heurística y evolutiva para la colocación de anclajes UWB.....	73
Matei ANDREI MADALIN and Lenuta ALBOAIE	
Evaluación de la calidad de imagen en imágenes de fantoma mamográfico.....	79
Valentina CORCHUELO, Juan MANUEL NÚÑEZ, Edison SALAZAR HURTADO y Hernán Darío BENÍTEZ RESTREPO	

Inteligencia artificial y aprendizaje federado para la gestión inteligente de huertos urbanos.... Johanna MORENO-CASTELLANOS, Laura GRANDE-PÉREZ, Juan M. NÚÑEZ V. and Sara RODRÍGUEZ G.	85
<i>Matching</i> semántico para la revitalización rural: una propuesta tecnológica ante la despoblación y la exclusión juvenil..... Natalia L.-GALLEGO, Isabella PÉREZ-CAVIEDES y Danna C. SALAMANCA-VALENZUELA	91
Plataforma accesible para alertas tempranas hidrometeorológicas basada en pronósticos climáticos y umbrales locales de precipitación..... Camila A. CARDONA-ALZATE y David Alejandro CAJIAO-LAZT	97
<i>Temporal Fusion Transformers</i> para la predicción del término de facturación PVPC..... Víctor GARCÍA-CHICO, Manuel RAMOS-JIMÉNEZ, Lucía RIERA-PRESMANES, Roberto SALAMANCA-BARRIOS y Ana Belén GIL GONZÁLEZ	103
Predicción del riesgo de incendios forestales mediante teledetección e IA (2020-2025)..... Rodrigo ÁLVAREZ, Javier SANTOS, Thide E. LLORENTE, Stephany NIEVES y Samuel REY	109
SEIR-XAI: inferencia explicable de la transmisión epidémica para análisis de escenarios y apoyo a decisiones..... Marco HERNÁNDEZ PÉREZ, Pablo-Enrique GUILLEM FERNÁNDEZ y Javier PRIETO	115
Sistema de aprendizaje adaptativo con IA para universitarios..... Santiago MURGUEITIO MUÑOZ y Juan Manuel NÚÑEZ VELASCO	119
Sistemas de tutoría inteligente y narrativas digitales para preservar la diversidad biocultural de Ecuador..... Juan Pablo LÓPEZ GOYEZ, Alfonso GONZÁLEZ BRIONES y Jesús ARANGUREN	125
Uso de K-means como herramienta de IA para la prevención de riesgos y síntomas respiratorios en poblaciones trabajadoras vulnerables colombianas..... Yenny Andrea ROZO SILVA, Leidy Isabel CALDERÓN SIERRA, Ana DELGADO-GARCÍA, Juan José AGÚN-GONZÁLEZ y Raúl AGUILAR ELENA	131
1.2. PITCH DE PROYECTOS	137
Adaptación, trazabilidad ética y cumplimiento normativo en agentes de inteligencia artificial conversacional: un marco de investigación inclusivo para la diversidad cognitiva y sensorial..... Teresa CELESTINO-BUENTELLO y Guillermo HERNÁNDEZ	139
Edificios inteligentes en entornos rurales: integración de IoT e inteligencia artificial en la provincia de Valladolid..... David ESCUDERO-MANCEBO y Jezabel GÓMEZ	145
Grandes modelos de lenguaje como mediador lingüístico en la atención médica..... Thalia RUIZ DESDÍN	151
Desafíos de la IA generativa en dominios de alto riesgo: requisitos para la simplificación judicial fiable..... Ana-Belén GIL-GONZÁLEZ, Silvia DOMÍNGUEZ CASTRO, Clara RODRÍGUEZ MARCOS, Alberto JIMÉNEZ SÁNCHEZ, Javier MANSO CONDE y Jorge Enrique GONZÁLEZ GONZALO	155
Nuvora: un marco de IA fiable para la salud digital integrada y el bienestar de los pacientes..... Jaime SUÁREZ PÉREZ	163

INDEX

Introduction.....	11
Dra. Ana BELÉN GIL	
1. ORAL COMMUNICATIONS	13
1.1. SCIENTIFIC CONTRIBUTIONS.....	15
Application of Artificial Intelligence to Predict the Impact of Climate Change on Vine Phenology in Burgundy (France): Implications for Wine Sustainability.....	17
Sofania M. ROJAS-LANDACAY, Juan M. NÚÑEZ V., José-Abel FLORES and Fernando DE LA PRIETA	
Transformations in Higher Education: Redefining the Teaching Role in the Face of the Emergence of Generative Artificial Intelligence	23
Claudia JIMENO POSTIGO	
Construction and Validation of an Ontology for the Development of Artificial Intelligence Solutions aimed at the Demographic Challenge.....	29
Jorge DOMÍNGUEZ	
CytoLens: Preliminary Framework for Supporting Screening and Prioritisation of Findings in Digital Cervical Cytology Using Artificial Intelligence.....	35
Sandra MARCOS RECIO, Lautaro ROSSI LABIANCA, Ana-Belén GIL-GONZÁLEZ y Andrés CARDONA-MENDOZA	
DataMarket: Intelligent Analytics Application for Optimisation and Loyalty in Neighbourhood Minimarkets	41
Juan Manuel LÓPEZ, Juan Felipe ZAMBRANO, Juan CAMILO BURBANO, Andrés ENRÍQUEZ and Juan M. NÚÑEZ V.	
Intelligent Modelling for Sustainable Forest Biomass Management Using Geospatial AI...	49
Sarai COBAS MERINO, Darío LÓPEZ RAMOS, Jairo CALDERÓN MICAÑA and Ana-Belén GIL-GONZÁLEZ	
Artificial Intelligence (AI) in Assisting the Elderly in Rural Settings: A Review of the State of the Art	55
Laura GRANDE-PÉREZ, Johanna MORENO-CASTELLANOS and Ana Belén GIL-GONZÁLEZ	
Foundations and Design of an AI-Based Service for Financial and Pharmacological Well-Being	61
José ORTEGA-MOHEDANO and Jaime SUÁREZ PÉREZ	
A Unified Tool for Studying Forest Crops in Spain.....	67
Jaime GABRIEL VEGAS, Maider Araceli URBON JIMÉNEZ , Ana DE LUIS REBOREDO, Belén PÉREZ LANCHO and Ana-Belén GIL-GONZÁLEZ	
Heuristic and Evolutionary Optimization for UWB Anchor Placement	73
Matei ANDREI MADALIN and Lenuta ALBOAIE	
Image Quality Assessment of Mammography Phantom Images.....	79
Valentina CORCHUELO, Juan MANUEL NÚÑEZ, Edison SALAZAR HURTADO and Hernán Darío BENÍTEZ RESTREPO	

Artificial Intelligence and Federated Learning for the Smart Management of Urban Gardens Johanna MORENO-CASTELLANOS, Laura GRANDE-PÉREZ, Juan M. NÚÑEZ V. and Sara RODRÍGUEZ G.	85
Semantic Matching for Rural Revitalization: A Technological Proposal to Address Depo- pulation and Youth Exclusion Natalia L.-GALLEGO, Isabella PÉREZ-CAVIEDES and Danna C. SALAMANCA-VALENZUELA	91
Accessible Platform for Hydrometeorological Early Warnings Based on Climate Forecasts and Local Precipitation Thresholds Camila A. CARDONA-ALZATE and David Alejandro CAJIAO-LAZT	97
Temporal Fusion Transformers (TFT) for the Prediction of the PVPC Billing Term V́ctor GARCÍA-CHICO, Manuel RAMOS-JIMÉNEZ, Lucía RIERA-PRESMANES, Roberto SALAMANCA-BARRIOS, and Ana Belén GIL GONZÁLEZ	103
Forest Fire Risk Prediction Using Remote Sensing and Artificial Intelligence Rodrigo ÁLVAREZ, Javier SANTOS, Thide E. LLORENTE, Stephany NIEVES and Samuel REY	109
SEIR-XAI: Explainable Inference of Epidemic Transmission for Scenario Analysis and De- cision Support..... Marco HERNÁNDEZ PÉREZ, Pablo-Enrique GUILLEM FERNÁNDEZy Javier PRIETO	115
AI-Based Adaptive Learning System for University Students Santiago MURGUEITIO MUÑOZ and Juan Manuel NÚÑEZ VELASCO	119
Intelligent Tutoring Systems and Digital Narratives to Preserve Ecuador's Biocultural Diversity ... Juan Pablo LÓPEZ GOYEZ, Alfonso GONZÁLEZ BRIONES and Jesús ARANGUREN	125
Use of K-means as an AI Tool for the Prevention of Respiratory Risks and Symptoms in Vulnerable Working Populations in Colombia Yenny Andrea ROZO SILVA, Leidy Isabel CALDERÓN SIERRA, Ana DELGADO-GARCÍA, Juan José AGÚN-GONZÁLEZ and Raúl AGUILAR ELENA	131
1.2. PITCH DE PROYECTOS Adaptation, Ethical Traceability and Regulatory Compliance in Conversational Artificial In- telligence Agents: An Inclusive Research Framework for Cognitive and Sensory Diversity..... Teresa CELESTINO-BUENTELLO and Guillermo HERNÁNDEZ	137
Smart Building in Rural Areas: Integration of IoT and Artificial Intelligence in the Province of Valladolid David ESCUDERO-MANCEBO and Jezabel GÓMEZ	145
Large Language Models as Language Mediators in Healthcare Thalia RUIZ DESDÍN	151
Challenges of Generative AI in High-Risk Domains: Requirements for Reliable Judicial Simplification Ana-Belén GIL-GONZÁLEZ, Silvia DOMÍNGUEZ CASTRO, Clara RODRÍGUEZ MARCOS, Alberto JIMÉNEZ SÁNCHEZ, Javier MANSO CONDE and Jorge Enrique GONZÁLEZ GONZALO	155
Nuvora: A Trustworthy AI Framework for Integrated Digital Health and Patient Well-Being... Jaime SUÁREZ PÉREZ	163

INTRODUCCIÓN

INTRODUCTION

Dra. Ana BELÉN GIL

Directora de la Cátedra Internacional en Inteligencia Artificial y Reto Demográfico (DemIA)
Universidad de Salamanca

 <https://orcid.org/0000-0001-7235-6151>

Esta publicación recopila las contribuciones científicas y las reflexiones académicas presentadas en el *primer simposio en Inteligencia artificial aplicada al reto demográfico*, una iniciativa impulsada por la *Cátedra Internacional en Inteligencia Artificial y Reto Demográfico (DemIA)*. Este encuentro, celebrado el 4 de diciembre de 2025 en la histórica ciudad de Béjar (Salamanca), nació con el propósito de crear un foro de referencia para abordar uno de los desafíos más urgentes de nuestra sociedad contemporánea: la convergencia entre la innovación tecnológica y la cohesión territorial.

Bajo el lema *IA fiable al servicio del reto demográfico*, el simposio se articuló en torno a la convicción de que la inteligencia artificial no debe ser un fin en sí misma, sino una herramienta ética, responsable y eficaz al servicio del bien común. En un contexto marcado por la despoblación, el envejecimiento y la necesidad de crear nuevas oportunidades en entornos rurales, la Cátedra DemIA organizó este encuentro para analizar cómo la tecnología puede contribuir al equilibrio territorial y a la justicia social.

El respaldo institucional fue clave para el éxito de la jornada, lo que puso de relieve la importancia de la colaboración internacional y multidisciplinar. Contamos con el apoyo fundamental de la *Escuela Técnica Superior de Ingeniería Industrial de Béjar (ETSII)* y su director, Alejandro Reveriego Martín, así como con la destacada participación de nuestros socios de la *Universitatea română Alexandru Ioan Cuza de Iași (UAIC)*, representados por su vicerrector de Digitalización, Adrian Iftene, y la decana de la Facultad de Ciencias Computacionales, Lenuta Alboai. Esta alianza propiciada por DemIA simboliza la escala global de un debate que requiere diversas perspectivas para diseñar una “IA fiable” basada en la transparencia, la gobernanza y, sobre todo, la confianza social.

El programa del simposio se estructuró en una intensa jornada que combinó ponencias magistrales, mesas redondas sobre transferencia tecnológica y sesiones técnicas en las que se presentaron un total de 27 contribuciones seleccionadas. La naturaleza híbrida del evento

permitió una participación amplia y diversa, que se materializó en 23 presentaciones orales y 4 pósteres. Las temáticas abordadas pusieron de manifiesto la transversalidad del desafío demográfico, siendo el área de IA y Salud la que más peso tuvo, con 10 contribuciones, seguida por las investigaciones en Agricultura y Transformación Digital, con 4 contribuciones cada una. El programa se completó con aportaciones esenciales en Ética y Gobernanza, Educación, Modelado demográfico y Despoblación, ofreciendo así una visión completa de la intersección entre tecnología y territorio.

Más allá de las cifras, las investigaciones recogidas en estas actas demuestran que la IA no es un concepto abstracto de futuro, sino una realidad presente. Desde la mejora en el cuidado de las personas mayores hasta el diseño de servicios públicos eficientes y la gestión inteligente de recursos, las páginas de este volumen ofrecen soluciones tangibles orientadas a humanizar la tecnología.

Uno de los momentos más destacados del encuentro fue la entrega de los Premios DemIA, que galardonaron trabajos de alto impacto social y técnico en reconocimiento a la excelencia investigadora. El premio a la Mejor Contribución Científica recayó en Matei Madalin (UAIC), por su investigación titulada «Heuristic and Evolutionary Optimization for UWB Anchor Placement in Indoor Environments», un avance clave para la localización precisa en interiores. Asimismo, Thalía Ruiz fue distinguida con el premio al Mejor Pitch por su propuesta «Grandes modelos de lenguaje como mediador lingüístico en la atención médica», que explora el uso de la IA para romper barreras comunicativas en el ámbito sanitario. Por último, Verónica Bolón Canedo recibió el reconocimiento en la categoría de Mejor TFM/TFG por su trabajo «Desarrollo de un modelo computacional para predecir la toxicidad de nanopartículas mediante aprendizaje automático», completando un cuadro de honor que evidencia el talento joven dedicado a resolver problemas reales.

El espíritu colaborativo que impregnó el encuentro en Béjar se refleja en la calidad de las contribuciones aquí reunidas. Estas investigaciones no solo aportan conocimiento técnico, sino que también sientan las bases para políticas públicas más inclusivas y nuevas líneas de investigación que trascienden el propio simposio.

Más que un punto de llegada, este primer simposio se ha consolidado como un espacio de apertura y el reflejo del diálogo permanente al que DemIA desea dar continuidad con un simposio anual. Estas actas son el testimonio del compromiso de la Cátedra DemIA y de todos los investigadores participantes de seguir trabajando en una inteligencia artificial que respete la diversidad de nuestros territorios y comunidades, y que garantice que el progreso tecnológico sea, ante todo, un progreso para las personas.

1

COMUNICACIONES ORALES

1.1.
CONTRIBUCIONES CIENTÍFICAS

APLICACIÓN DE INTELIGENCIA ARTIFICIAL PARA PREDECIR EL IMPACTO DEL CAMBIO CLIMÁTICO EN LA FENOLOGÍA DE LA VID EN BORGOÑA (FRANCIA): IMPLICACIONES PARA LA SOSTENIBILIDAD VITIVINÍCOLA

APPLICATION OF ARTIFICIAL INTELLIGENCE TO PREDICT THE IMPACT OF CLIMATE CHANGE ON VINE PHENOLOGY IN BURGUNDY (FRANCE): IMPLICATIONS FOR WINE SUSTAINABILITY

Sofania M. ROJAS-LANDACAY¹, Juan M. NÚÑEZ V.², José-Abel FLORES³
y Fernando DE LA PRIETA⁴

¹ *Universidad de Salamanca. Salamanca, Castille and León, España*

✉ sofania@usal.es

🌐 <https://orcid.org/0009-0009-1788-9233>

² *Universidad de Salamanca. Salamanca, Castille and León, España*

✉ jmnunez@usal.es

🌐 <https://orcid.org/0000-0002-3165-6200>

³ *Universidad de Salamanca. Salamanca, Castille and León, España*

✉ flores@usal.es

🌐 <https://orcid.org/0000-0003-1909-293X>

⁴ *Universidad de Salamanca. Salamanca, Castille and León, España*

✉ fer@usal.es

🌐 <https://orcid.org/0000-0002-8239-5020>

RESUMEN: El cambio climático ha afectado significativamente a las etapas fenológicas, adelantando especialmente las etapas que influyen directamente en la calidad y la composición de la uva. Este estudio evaluó cómo el cambio climático influyó en las fases fenológicas de la viticultura de Bourgogne, en parcelas de Pinot Noir a través de la comparación de datos climáticos históricos (1966-2023) con información fenológica. Empleamos modelos de regresión por mínimos cuadrados ordinarios (OLS) y modelos avanzados de inteligencia artificial para modelar las relaciones no lineales entre las variables climáticas y las etapas fenológicas. Estos resultados subrayan el potencial de los modelos basados en IA como herramienta robusta y

escalable para la viticultura de precisión, la adaptación al cambio climático y el mantenimiento de la actividad vitivinícola como motor socioeconómico en territorios rurales.

Palabras clave: inteligencia artificial; cambio climático; viticultura.

ABSTRACT: Climate change has significantly affected phenological stages, particularly advancing those stages that directly influence grape quality and composition. This study evaluated how climate change influenced the phenological phases of viticulture in Burgundy, on Pinot Noir plots. This was done by comparing historical climate data (1966-2023) with phenological information. We used ordinary least squares (OLS) regression models and advanced artificial intelligence models to model the non-linear relationships between climate variables and phenological stages. These results highlight the potential of AI-based models as a robust and scalable tool for precision viticulture, climate change adaptation, and the maintenance of wine production as a socio-economic driver in rural areas.

Keywords: artificial intelligence; climate change; viticulture.

1. INTRODUCCIÓN

El cambio climático está transformando el panorama de la viticultura mundial, con efectos especialmente notorios en las regiones del Mediterráneo europeo, como España, Francia e Italia. Estas áreas, que históricamente han sido propicias para la producción vitivinícola de alta calidad, enfrentan un aumento sostenido de las temperaturas y una creciente aridez, factores que amenazan la viabilidad del cultivo de la vid. Se ha estimado que, en un futuro cercano, algunas de estas zonas podrían volverse inadecuadas para la producción, obligando a los viticultores a adaptar sus estrategias o incluso a migrar a regiones con condiciones más favorables [1, 2].

Estos cambios no solo afectarían la producción de vino en términos de volumen y calidad, sino que también tendrían un impacto socioeconómico en las regiones cuya economía depende en gran medida de esta actividad. Ante la creciente incertidumbre climática y sus efectos en la producción vitivinícola, se torna crucial el desarrollo de herramientas predictivas que permitan a los productores anticipar cambios fenológicos, facilitando la adaptación y la optimización de las estrategias de manejo. En este contexto, la aplicación de técnicas avanzadas de inteligencia artificial se presenta como un enfoque altamente prometedor para incrementar la precisión de las predicciones y mejorar la gestión vitivinícola en escenarios de cambio climático [2, 3, 4].

Los modelos predictivos aplicados en viticultura, en particular aquellos que integran inteligencia artificial (IA), han demostrado un alto potencial para mejorar la precisión de las predicciones fenológicas y de rendimiento de la uva, al aprovechar datos climáticos, fenológicos y geoespaciales [5]. Varias aplicaciones impulsadas por la IA en la viticultura, incluidos los modelos de predicción y los enfoques basados en datos para optimizar la gestión del agua y los nutrientes, están impulsando colectivamente la transición hacia

prácticas más sostenibles en el sector vitivinícola, permitiendo avanzar en la predicción de calidad de la uva y la respuesta de la vid al cambio climático [6, 7].

Frente a esto, el objetivo del estudio consistió en evaluar cómo el cambio climático influyó en las fases fenológicas de la viticultura en la región de Bourgogne, específicamente en las parcelas de *Pinot Noir* en la Côte de Beaune.

2. METODOLOGÍA

El estudio se llevó a cabo en la Côte de Beaune (47,0244° N, 4,8410° E), en el este de Borgoña (Francia), una región vinícola reconocida por su diversidad topográfica y climática que influyen directamente en la fenología de variedades como *Pinot Noir* [8, 9]. Los datos fenológicos de *Pinot Noir* (brotamiento, floración, envero y vendimia) fueron proporcionados por el BIVB (Bureau Interprofessionnel des Vins de Bourgogne). Los datos climáticos diarios (temperatura y precipitaciones) para el periodo 1966-2023 se obtuvieron de la estación de Dijon-Longvic a través del portal API de Météo-France. Los análisis se realizaron en Python y se evaluaron utilizando R^2 y RMSE. Las variables dependientes fueron los eventos fenológicos, mientras que las variables independientes fueron las variables bioclimáticas. El modelo OLS sigue la metodología descrita en [8], mientras que la aplicación de Random Forest se basa en los estudios de [10] y [11].

3. RESULTADOS

Los resultados revelaron que, debido al aumento de las temperaturas, las etapas fenológicas se han adelantado en promedio en las últimas décadas, confirmando que el cambio climático está alterando significativamente la fenología de la vid. Esta aceleración de las etapas fenológicas afecta directamente a la calidad y la producción del vino. Los gráficos generados por el análisis de regresión lineal (OLS) y Random Forest reflejan la evolución de las variables meteorológicas de temperatura y precipitación, mostrando esta tendencia. En particular, las temperaturas máximas (TX) muestran un aumento significativo a partir de 1970, lo que indica un cambio hacia un clima más cálido. Del mismo modo, se observa que las temperaturas mínimas [Figura 1 (a) y (b)] han experimentado un aumento progresivo, lo que sugiere una disminución de la amplitud térmica y favorece una maduración más rápida de la uva.

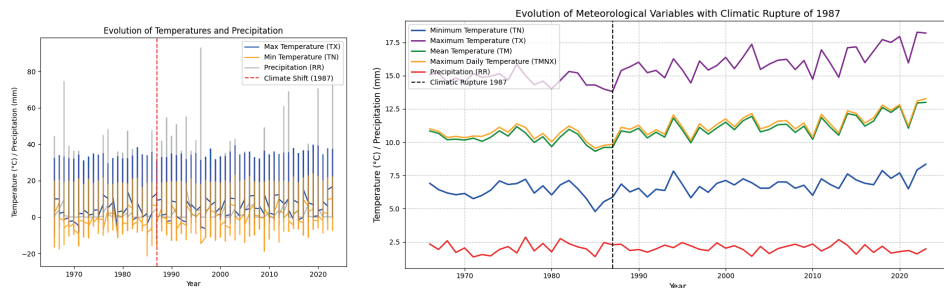


Figura 1. (a) y (b) Evolución de las variables meteorológicas con la ruptura climática de 1987 en la zona de Côte de Beaune (1966-2023)

Los resultados de este estudio validan la hipótesis de que el cambio climático, al aumentar las temperaturas, tiene un impacto directo en la fenología de la vid, acelerando las etapas fenológicas [12, 13]. Este fenómeno demuestra que las temperaturas más altas favorecen la acumulación de calor durante la temporada de crecimiento, acelerando la maduración de la uva y adelantando la cosecha [7]. El aumento implica un mayor estrés térmico en las vides, lo que afecta a su composición química, especialmente al equilibrio de azúcares, acidez y compuestos fenólicos, lo que podría alterar las características organolépticas del vino producido en la región [14].

En este estudio, el modelo permitió identificar variables climáticas clave, como la temperatura máxima con influencia directa sobre las fases fenológicas. Esta capacidad predictiva, junto con la posibilidad de escalar el modelo a otras regiones o condiciones climáticas futuras, convierte a la IA en una herramienta estratégica para anticipar impactos y diseñar respuestas adaptativas en viticultura de precisión. La integración de inteligencia artificial en la viticultura no solo permite mejorar la producción, sino que también ofrece soluciones para la revitalización de zonas rurales, asegurando que las comunidades agrícolas puedan adaptarse a los cambios climáticos y seguir siendo viables desde el punto de vista económico y social. Este tipo de estudios permite que los viñedos continúen siendo una actividad productiva viable en áreas rurales, contribuyendo a fijar población y fomentar el empleo local.

4. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation). Adicionalmente, queremos agradecer a BIVB por proporcionarnos

13. Alikadic, A., Pertot, I., Eccel, E., Dolci, C., Zarbo, C., Caffarra, A., de Filippi, R., & Furlanello, C. (2019). The impact of climate change on grapevine phenology and the influence of altitude: A regional study. *Agricultural and Forest Meteorology*, 271, 73-82. <https://doi.org/10.1016/j.agrformet.2019.02.030>
14. Tonietto, J., & Carbonneau, A. (2004). A multicriteria climatic classification system for grape-growing regions worldwide. *Agricultural and Forest Meteorology*, 124(1-2), 81-97. <https://doi.org/10.1016/j.agrformet.2003.06.001>

TRANSFORMATIONS IN HIGHER EDUCATION: REDEFINING THE TEACHING ROLE IN THE FACE OF THE EMERGENCE OF GENERATIVE ARTIFICIAL INTELLIGENCE

TRANSFORMACIONES EN LA EDUCACIÓN SUPERIOR: REDEFINICIÓN DEL PAPEL DEL PROFESORADO ANTE LA APARICIÓN DE LA INTELIGENCIA ARTIFICIAL GENERATIVA

Claudia JIMENO POSTIGO

University of Salamanca. Salamanca, Castilla y León, Spain

✉ jimenopostigo.claudia@usal.es

🌐 <https://orcid.org/0009-0000-8671-3033>

ABSTRACT: Generative artificial intelligence (GenAI) is an emerging technology that has significantly transformed the interaction between humans and machines. GenAI has the capacity to create content such as text, images, and videos, and it even uses human language. In the educational field, tools such as ChatGPT stand out for their ability to maintain coherent conversations, simulating human interactions. This study aims to offer a comprehensive and critical view of the convergence of GenAI and higher education. To this end, a systematic literature review has been carried out following the PRISMA protocol through the WoS and Dialnet databases. The analysis focuses on understanding the role of GenAI in this context, identifying both the opportunities and challenges associated with its implementation. The results of the study highlight key challenge areas, promising trends, and future prospects. Likewise, the effects of GenAI on students and teachers are analyzed, paying special attention to the ethical and social implications that accompany its integration into higher education.

Keywords: artificial intelligence; higher education; pedagogical use; ChatGPT; educational tool; pedagogical innovation.

RESUMEN: La inteligencia artificial generativa (GenAI) es una tecnología emergente que ha transformado significativamente la interacción entre humanos y máquinas. La GenAI tiene la capacidad de crear contenidos como textos, imágenes y vídeos, e incluso utiliza el lenguaje humano. En el ámbito educativo, herramientas como ChatGPT destacan por su capacidad para mantener conversaciones coherentes, simulando interacciones humanas. Este estudio tiene como objetivo ofrecer una visión global y crítica de la convergencia de la GenAI y la educación.

superior. Para ello, se ha llevado a cabo una revisión sistemática de la literatura siguiendo el protocolo PRISMA a través de las bases de datos WoS y Dialnet. El análisis se centra en comprender el papel de la GenAI en este contexto, identificando tanto las oportunidades como los retos asociados a su implementación. Los resultados del estudio destacan las áreas clave de desafío, las tendencias prometedoras y las perspectivas de futuro. Asimismo, se analizan los efectos de la GenAI en los estudiantes y profesores, prestando especial atención a las implicaciones éticas y sociales que acompañan a su integración en la educación superior.

Palabras clave: inteligencia artificial; educación superior; uso pedagógico; ChatGPT; herramienta educativa; innovación pedagógica.

1. INTRODUCTION ABOUT AI INTEGRATION

1.1. DEFINITION OF ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) is a complex concept without a single universally accepted definition. The European Commission defines it as software (and possibly hardware) systems designed by humans that, given a complex objective, act in the physical or digital dimension by perceiving their environment, interpreting data, processing information, and deciding on the best actions to achieve predetermined goals [1]. For Bryson [2], AI refers to «artifacts used to detect contexts or carry out actions in response to those contexts that have already been detected» (pp. 127-159).

Other definitions describe AI as «the ability of a system to correctly interpret external data, learn from that data, and use that learning to achieve specific goals and tasks through flexible adaptation» [3] or as «any technique that enables computers to mimic human behavior» [4]. Ultimately, AI encompasses systems capable of performing tasks that typically require human intelligence, such as learning, reasoning, and perception.

1.2. THE ROLE OF THE TEACHER SINCE THE EMERGENCE OF AI

The arrival of generative AI tools like ChatGPT has significantly transformed the educator's role in higher education. Authors such as [5] posit that teachers now act more as facilitators of learning rather than as the primary source of knowledge.

- *From Transmitter to Facilitator and Mentor:* With AI, teachers can guide students to explore and understand information independently, fostering a more student-centered approach [6]. This allows educators to focus on developing critical skills like critical thinking and problem-solving.
- *Designer of Personalized Learning Experiences:* AI enables the creation of personalized and adaptive learning experiences. Teachers become designers of these experiences, using tools like ChatGPT to develop interactive activities tailored to individual student needs [7].

- *Continuous Training and Professional Development:* Effective integration requires ongoing teacher training. Educators need to learn to use these tools pedagogically, interpret the data they generate, and address associated ethical challenges [8].
- *Assessment and Feedback:* ChatGPT and similar tools transform assessment by enabling more timely and precise automated feedback, allowing teachers to dedicate more time to continuous and formative evaluation rather than traditional summative exams [9].

In summary, the teacher's role evolves towards educational guidance, pedagogical design, and ethical mediation in the use of AI, leveraging technology to enhance the teaching-learning process without replacing the essential human element.

1.3. GENERATIVE AI IN HIGHER EDUCATION

Generative Artificial Intelligence (GAI), a subfield of AI focused on autonomously creating content (text, images, etc.), has emerged as a disruptive force in higher education.

- *Potential for Transformation:* GAI has the potential to revolutionize teaching, learning, and research by automating content generation, personalizing instruction, and providing real-time data analysis [10].
- *Key Applications in the Academic Context:*
 1. *Personalized Learning:* Adaptive systems can tailor educational resources and pathways to individual student needs and learning styles.
 2. *Automated Tutoring and 24/7 Support:* Tools like ChatGPT can offer instant assistance, answer questions, and provide explanations outside of classroom hours.
 3. *Content and Resource Creation:* AI can help generate educational materials, draft summaries, or create practice exercises, saving educators time.
 4. *Data Analysis and Research Support:* GAI can assist in literature reviews, data processing, hypothesis generation, and identifying at-risk students through learning analytics [11, 12].
- *Inherent Challenges:* Despite its potential, GAI integration poses significant challenges:
 1. *Academic Integrity:* Risks of plagiarism, unauthorized use in assessments, and questions about authorship [13].
 2. *Information Reliability:* Potential for generating inaccurate or biased information («hallucinations»).
 3. *Ethical and Privacy Issues:* Concerns regarding data privacy, algorithmic bias, and equitable access to technology.
 4. *Need for New Competencies:* Highlights the urgent need to develop AI literacy among both students and teachers.

1.4. AI APPLICATIONS AND TRENDS IN EDUCATION

The application of AI in education (AIEd) is based on a wide range of techniques that go beyond GAI:

Table 1. Core technique underlying AI applications

Technique	Description	Application
<i>Learning Analytics & Educational Data Mining</i>	Analysis of large student datasets to identify patterns and trends.	Design of personalized curricula and provision of detailed feedback.
<i>Multi-Agent Systems</i>	Creation of intelligent tutoring systems that interact in real-time.	Personalized and adaptive learning experiences.
<i>Neural Networks</i>	Used for classification, data analysis, and pattern recognition.	Learning pattern identification, voice/text recognition systems.
<i>Fuzzy Logic</i>	Modeling of uncertainty and creation of adaptive systems.	Development of adaptive and personalized assessment systems.
<i>Bayesian Systems</i>	Decision-making and uncertainty modeling.	Adaptive assessment systems.

Source: Adapted from Jimeno (2024).

2. CURRENT AND EMERGING TRENDS

- *Conversational AI and Advanced Chatbots:* Evolution towards more specialized and contextual conversational agents, moving beyond general-purpose language models like ChatGPT to provide support and functionalities tailored to specific domains.
- *Human Capability Augmentation with AI:* Development of tools designed to enhance and complement human capabilities across various fields, automating routine tasks (such as data analysis or content generation) so professionals can focus on higher-value activities like strategic decision-making, creativity, and complex interaction.
- *Predictive Analytics:* Increasing application of AI algorithms to analyze patterns in large datasets, enabling the prediction of outcomes, behaviors, or needs, which facilitates proactive intervention and the personalization of services or products.
- *Ethics and Regulation:* Growing focus on the development of ethical frameworks, responsible use guidelines, and regulatory frameworks to ensure fairness, transparency, privacy, and safety in the design and deployment of AI systems, mitigating risks such as algorithmic bias.
- *AI Literacy as a Core Competency:* Recognition of the need to integrate skills for understanding, interacting with, and critically evaluating AI into training programs and professional development, preparing society for an environment increasingly mediated by this technology.

3. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

4. REFERENCES

1. Bryson, J. J. (2018). La última década y el futuro del impacto de la IA en la sociedad. *OpenMind*.
2. Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
3. Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, 2, Artículo 4. <https://doi.org/10.1007/s44163-022-00022-8>
4. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
5. Davenport, T. H., & Kirby, J. (2016). Only humans need apply: Winners and losers in the age of smart machines. Harper Business.
6. González-González, C. S. (2023). El impacto de la inteligencia artificial en la educación: Transformación de la forma de enseñar y de aprender. *Revista Currículum*, 36(3), 15-34. <https://doi.org/10.25145/j.qurricul.2023.36.03>
7. Gallent-Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: Una mirada desde la ética y la integridad académica. *RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa*, 29(2), Artículo 4. <https://doi.org/10.30827/relieve.v29i2.29134>

CONSTRUCCIÓN Y VALIDACIÓN DE UNA ONTOLOGÍA PARA EL DESARROLLO DE SOLUCIONES DE INTELIGENCIA ARTIFICIAL ORIENTADAS AL RETO DEMOGRÁFICO

CONSTRUCTION AND VALIDATION OF AN ONTOLOGY FOR THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE SOLUTIONS AIMED AT THE DEMOGRAPHIC CHALLENGE

Jorge DOMÍNGUEZ

Universidad de Salamanca. Salamanca, Castilla y León, España

✉ jorge.dommar@usal.es

🌐 <https://orcid.org/0009-0005-4665-5437>

RESUMEN: Este trabajo presenta la construcción y la validación de una ontología estructurada como base de conocimiento para el desarrollo de soluciones de IA orientadas al reto demográfico. El proyecto aborda dominios clave como suelos, Industria 4.0, servicios inteligentes, entre otros.

Palabras clave: ontología; base de conocimiento; reto demográfico; territorio inteligente; GeoSPARQL; Industria 4.0.

ABSTRACT: This work presents the construction and validation of a structured ontology as a knowledge base for the development of AI solutions aimed at addressing demographic challenges. The project addresses key domains such as land, Industry 4.0, and smart services, among others.

Keywords: ontology; knowledge base; demographic challenge; smart territory; GeoSPARQL; Industry 4.0.

1. INTRODUCCIÓN

La gestión eficiente de los territorios rurales y la lucha contra la despoblación requieren superar la fragmentación de datos existentes entre sectores estratégicos. Las soluciones de

software tradicionales tienden a crear «silos de datos» que impiden una visión integral del territorio [1], dificultando la implementación de políticas públicas efectivas como las relacionadas con la PAC o los fondos *NextGen*. En este contexto, la adopción de tecnologías de la web semántica y el modelado ontológico se presenta como una solución robusta para garantizar la interoperabilidad y la representación formal del conocimiento [2].

El concepto de *Smart Village* o territorio inteligente implica la integración de dimensiones físicas, sociales y digitales [3]. Sin embargo, la mera digitalización de activos no es suficiente si no existe una capa semántica que conecte, por ejemplo, la producción agrícola con la demanda energética industrial o la oferta turística con la revitalización demográfica.

Este proyecto describe el desarrollo de una «ontología de integración» que actúa como un *hub* semántico, unificando cuatro ontologías de dominio vertical previamente desarrolladas (Agricultura, Industria, Despoblación y Turismo). El objetivo no es crear un mero contenedor, sino una capa de orquestación capaz de modelar reglas de negocio transversales y responder a preguntas complejas que ninguna ontología sectorial podría resolver por sí sola.

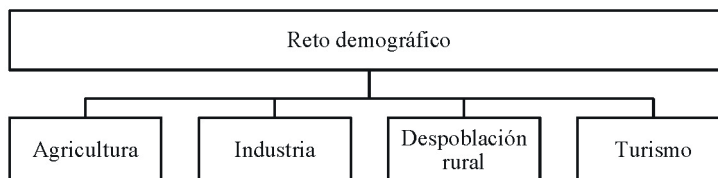
2. METODOLOGÍA Y ESTÁNDARES

Para abordar la complejidad inherente al reto demográfico se ha descartado el desarrollo de una ontología monolítica en favor de una arquitectura modular y federada. Este enfoque facilita la reutilización, la gestión de la complejidad y el mantenimiento descentralizado del conocimiento [4].

La metodología de construcción se ha alineado con estándares consolidados para asegurar la interoperabilidad industrial europea. El desarrollo extiende el ecosistema SAREF (*Smart Appliances REFerence ontology*) del ETSI, integrando extensiones específicas como SAREF4Agri o SAREF4City y complementándose con estándares geoespaciales como GeoSPARQL para la localización precisa de activos en el territorio.

El ciclo de vida del desarrollo ontológico ha seguido una metodología iterativa basada en las prácticas de ingeniería ontológica establecidas [5, 6], utilizando *Turtle* (.ttl) como formato nativo por su compatibilidad con los estándares W3C (RDF/OWL) y herramientas de validación lógica formal como el razonador *HermiT*.

Figura 1. Arquitectura jerárquica del enfoque modular federado



3. ONTOLOGÍA DE INTEGRACIÓN

Esta ontología de integración constituye el núcleo de la propuesta. Su función principal es modelar las interdependencias complejas de un territorio real, donde las decisiones en un sector repercuten inevitablemente en otros.

3.1. REGLAS DE NEGOCIO TRANSVERSALES

La integración se materializa mediante la definición de clases y propiedades «puente» que conectan dominios disjuntos provenientes de los módulos verticales. Se han implementado tres reglas de negocio principales que sustentan la lógica del grafo: economía circular (flujo de valorización de residuos), sinergia turismo-sociedad (importancia de activos turísticos) y cadena de valor agroindustrial (trazabilidad completa desde la producción en el sector primario hasta la industria de transformación).

3.2. ANÁLISIS MÉTRICO Y ONTOLOGÍA RESULTANTE

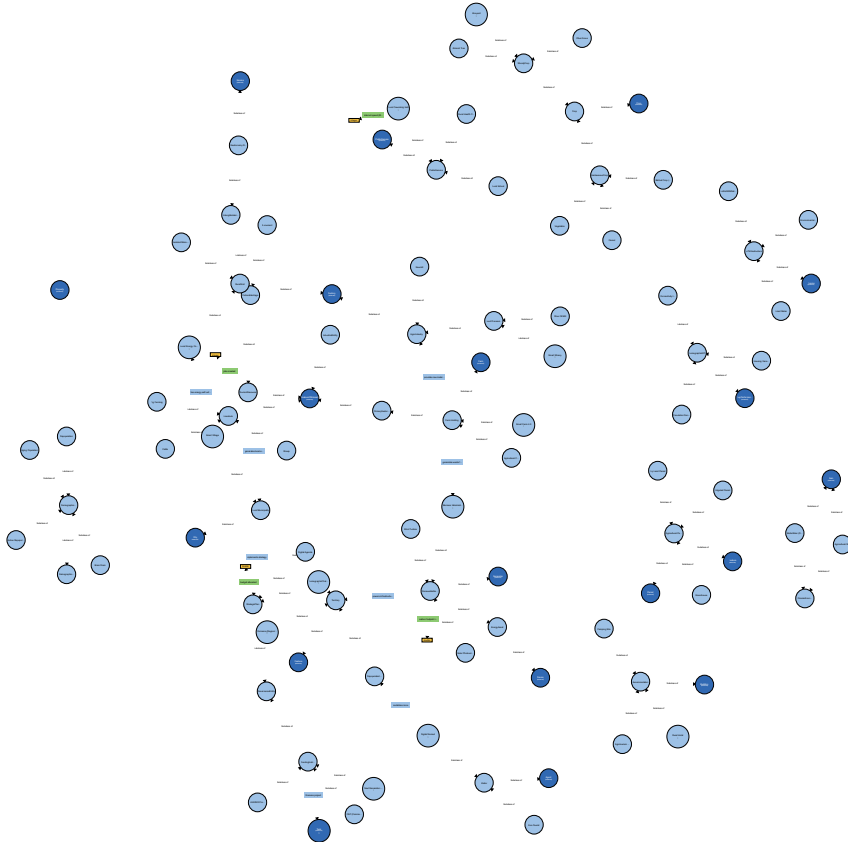
La robustez de esta ontología se evidencia en sus métricas estructurales, las cuales reflejan la suma del conocimiento federado y la complejidad de las reglas de integración. El análisis cuantitativo realizado con la herramienta de construcción (*Protégé*) arroja los siguientes resultados:

Tabla 1. Métricas de la ontología de integración

Métrica	Valor	Interpretación
Número de clases	100	Conceptos que cubren el espectro completo del territorio rural.
Propiedades de objeto	8	Relaciones «puente» de alto nivel definidas anteriormente que conectan dominios disjuntos. Permiten la trazabilidad circular entre sectores.
Propiedades de datos	4	Definidas específicamente para la capa de gestión y gobernanza (e. g. presupuestos de estrategias o KPIs agregados), necesarias para evaluar el escenario global.
Individuos	12	Instancias que verifican el escenario expuesto.
Axiomas lógicos	135	Indicador de densidad de reglas lógicas, un valor elevado indica un gran número de restricciones y definiciones formales utilizadas para inferir conocimiento.
Total de axiomas	316	Carga total de información de la ontología. Una magnitud elevada evidencia un alto grado de documentación y formalización.

El elevado número de clases y axiomas lógicos indican un alto grado de cobertura del espectro completo del territorio rural, esencial para inferir nuevo conocimiento automáticamente [7].

Figura 2. Grafo vectorizado de la ontología de integración generado con *Widoco*



4. CONCLUSIONES

La construcción de esta ontología representa un avance significativo en la aplicación de inteligencia artificial para el reto demográfico. Al federar el conocimiento de dominios clave bajo estándares europeos, se ha logrado pasar de silos de información aislados a un grafo de conocimiento conectado y razonable, sirviendo como base sólida para futuros sistemas de soporte a la decisión en la Administración Pública y el sector privado.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

6. REFERENCIAS

1. He, X., Kuai, X., Li, X., Qiu, Z., He, B., & Guo, R. (2025). Smart city ontology framework for urban data integration and application. *Smart Cities*, 8, Artículo 165. <https://doi.org/10.3390/smartcities8050165>
2. Pinto, A., Cardinale, Y., Dongo, I., & Ticona-Herrera, R. (2022). An ontology for modeling cultural heritage knowledge in urban tourism. *IEEE Access*, 10, 61820-61842. <https://doi.org/10.1109/ACCESS.2022.3179664>
3. Żukowska, S., Chmiel, B., & Połom, M. (2023). The smart village concept and transport exclusion of rural areas: A case study of a village in northern Poland. *Land*, 12(1), Artículo 260. <https://doi.org/10.3390/land12010260>
4. Huang, Y., Dhoub, S., Medinacelli, L. P., & Malenfant, J. (2023). Semantic interoperability of digital twins: Ontology-based capability checking in AAS modeling framework. In *Proceedings - 2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems (ICPS)*. IEEE. <https://doi.org/10.1109/ICPS58381.2023.10128003>
5. Vallbé, J.-J., & Barrera, M. F. (s.f.). *Lecturas sobre derecho y web semántica*.
6. Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology* (Stanford Knowledge Systems Laboratory Technical Report KSL-01-05). Stanford University.
7. Kakderi, C., & Panori, A. (2021). *The smart city ontology 2.0: Assessing the components and interdependencies of city smartness* (Preprint). Preprints. <https://doi.org/10.20944/preprints202108.0101.v1>

CYTOLENS: MARCO PRELIMINAR PARA EL APOYO AL CRIBADO Y LA PRIORIZACIÓN DE HALLAZGOS EN CITOLOGÍA CERVICAL DIGITAL MEDIANTE INTELIGENCIA ARTIFICIAL

CYTOLENS: PRELIMINARY FRAMEWORK FOR SUPPORTING SCREENING AND PRIORITISATION OF FINDINGS IN DIGITAL CERVICAL CYTOLOGY USING ARTIFICIAL INTELLIGENCE

Sandra MARCOS RECIO¹, Lautaro ROSSI LABIANCA², Ana-Belén GIL-GONZÁLEZ³
y Andrés CARDONA-MENDOZA⁴

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ Sandra_Marcos@usal.es

🌐 <https://orcid.org/0009-0009-3379-7304>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ lross98@usal.es

🌐 <https://orcid.org/0009-0000-7562-7586>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ abg@usal.es

🌐 <https://orcid.org/0000-0001-7235-6151>

⁴ *Universidad El Bosque, Bogotá, Colombia*

✉ acardonam@unbosque.edu.co

🌐 <https://orcid.org/0000-0002-6697-5471>

RESUMEN: El proyecto CytoLens, en colaboración entre la Universidad El Bosque y BISITE, propone un marco metodológico para apoyar el cribado citológico del cáncer cervical en patología digital mediante inteligencia artificial a partir de láminas completas digitalizadas (WSI). El trabajo define un pipeline reproducible de control de calidad y generación de parches, y en su estado actual se centra en detectores de instancias tipo YOLO para aportar evidencia localizada y soporte a la revisión experta. Como evolución del marco, se contempla incorporar modelos de aprendizaje por instancias múltiples (MIL) con autoatención para priorización a nivel de lámina y estrategias de pseudoetiquetado con revisión parcial, con evaluación basada en particionado por paciente y criterios de utilidad clínica.

Palabras clave: citología cervical; inteligencia artificial; whole slide imaging; detección de instancias; multiple instance learning; Bethesda; YOLO.

ABSTRACT: The CytoLens project, a collaboration between El Bosque University and BISITE, proposes a methodological framework to support the cytological screening of cervical cancer in digital pathology using artificial intelligence based on whole-slide images (WSIs). The work defines a reproducible pipeline for quality control and patch generation, and in its current state focuses on YOLO-type instance detectors to provide localised evidence and support for expert review. As a further development of the framework, the plan is to incorporate multi-instance learning (MIL) models with self-attention for slide-level prioritisation and pseudo-labelling strategies with partial review, with evaluation based on patient-level partitioning and clinical utility criteria.

Keywords: cervical cytology; artificial intelligence; whole slide imaging; instance detection; multiple instance learning; Bethesda; YOLO.

1. INTRODUCCIÓN

El cáncer de cuello uterino es prevenible en gran medida mediante programas de cribado que permitan identificar y tratar lesiones precursoras. No obstante, la citología cervical sigue requiriendo inspección visual sistemática de muestras con alta variabilidad de tinción y calidad, con artefactos, solapamientos y baja celularidad, lo que incrementa la incertidumbre diagnóstica y la variabilidad interobservador [1-3]. En consecuencia, los sistemas de apoyo deben formularse explícitamente como herramientas de triaje (priorización), donde el objetivo operativo es sostener alta sensibilidad en escenarios con desbalance marcado y con restricciones de tiempo de lectura por lámina [1-3].

La digitalización completa de láminas (WSI) traslada el proceso a un entorno computacional donde es posible localizar regiones informativas, estimar riesgo a nivel de lámina y generar evidencia visual para revisión rápida y consistente [3-5]. En este escenario, CytoLens se formula como herramienta asistencial orientada a priorización de casos y señalización de regiones/células sospechosas, con salidas compatibles con Bethesda [6, 7]. Este trabajo describe un marco jerárquico (lámina→región→célula) y un protocolo reproducible de extracción/selección de regiones para su evaluación posterior.

2. TRABAJOS RELACIONADOS Y MOTIVACIÓN

Los enfoques de aprendizaje profundo en citología han evolucionado desde la clasificación de células aisladas hacia la detección de instancias y el análisis sobre WSI completas. Detectores de una etapa tipo YOLO han mostrado capacidad para localizar células en escenas densas y con solapamiento, con tiempos de inferencia compatibles con el flujo de laboratorio [8]. Sin embargo, su escalabilidad se ve limitada por el coste de anotación

exhaustiva y por la pérdida de rendimiento ante cambios de centro, escáner o protocolo [1-3]. Para mitigar el cambio de dominio (diferencias entre laboratorios, escáneres, protocolos y tinciones), el diseño experimental debe incorporar análisis de sensibilidad a variaciones de tinción y desenfoque mediante aumentos controlados y, cuando sea posible, validación externa o evaluación entre centros [1-3].

Para reducir dependencia de anotaciones finas, la arquitectura multiple instance learning (MIL) permite entrenar con etiqueta global por lámina y aprender a ponderar parches informativos mediante mecanismos de agregación/atención [9]. En patología digital, enfoques a nivel de lámina basados en agregación de parches/detecciones han permitido modelos de triaje con menor granularidad de anotación [5, 10], mientras que MIL formaliza este aprendizaje bajo supervisión débil [9]. No obstante, la aceptación clínica exige enlazar la predicción global con evidencia local interpretable. Las revisiones recientes subrayan como retos dominantes la variabilidad interlaboratorio, el desbalance de clases y la necesidad de validación externa y trazabilidad [1-3].

3. MARCO CYTOLENS

3.1. ARQUITECTURA JERÁRQUICA

CytoLens se concibe como un pipeline modular para citología cervical digital sobre WSI. En una primera etapa, la WSI se divide en parches tras filtrado de fondo y normalización de color, con el objetivo de construir conjuntos de datos consistentes para entrenamiento y evaluación. En el estado actual del proyecto, el foco se sitúa en detectores de instancias basados en YOLO, entrenados sobre parches para localizar y categorizar células o agrupaciones sospechosas, generando evidencia visual para revisión por especialistas [8].

Como línea de evolución, se plantea incorporar un módulo de multiple instance learning (MIL) para trabajar con supervisión más débil a nivel de lámina y priorizar regiones candidatas mediante mecanismos de agregación/atención; en ese esquema, la salida de MIL guiaría la selección de regiones y YOLO aportaría evidencia localizada sobre las zonas priorizadas [8, 9]. Estas extensiones (MIL y pseudoetiquetado para ampliar datos con revisión parcial) se describen como componentes planificados y no como resultados ya consolidados, manteniendo el enfoque en trazabilidad e interpretabilidad [1-3].

3.2. DISEÑO EXPERIMENTAL Y EVALUACIÓN

El protocolo experimental prioriza el particionado por paciente para evitar fugas de información y estimaciones optimistas. Se contemplan escenarios uniclase y multiclase, con agregación operativa de categorías Bethesda cuando la disponibilidad de etiquetas no cubra toda la taxonomía [6, 7]. Para los detectores de instancias, se prevé emplear métricas estándar de detección (p. ej., mAP a umbrales IoU definidos) y análisis estratificados por tipo de lesión o subconjuntos celulares; adicionalmente, se plantea reportar métricas de

cribado a nivel de caso/lámina cuando el diseño experimental lo permita (p. ej., sensibilidad/especificidad en esquemas de triaje) [1-3, 8].

En línea con los retos descritos en revisiones recientes, el diseño incluye análisis de robustez ante variabilidad de tinción, artefactos y cambios de adquisición, mediante configuraciones controladas de aumento de datos y comparativas por subconjuntos [1-3]. El diseño incorpora trazabilidad de datos, control de versiones del pipeline y registro de configuraciones experimentales, en coherencia con principios de buenas prácticas de ML en dispositivos médicos [11].

4. ESTADO ACTUAL Y PRÓXIMOS PASOS

Actualmente se dispone de un pipeline reproducible para la extracción de parches desde WSI y la construcción de conjuntos de datos de detección a resoluciones controladas. Sobre esta base, se han configurado y entrenado múltiples variantes de modelos YOLO (p. ej., diferentes tamaños/capacidades) para caracterizar el compromiso velocidad-rendimiento y la estabilidad frente a aumentos orientados a variaciones de tinción y desenfoque, como paso previo a la ampliación de cohortes y a la maduración del protocolo de evaluación [8].

Los siguientes pasos incluyen: (i) consolidación del diseño experimental con particionado estricto por paciente y análisis por subgrupos; (ii) ampliación progresiva del conjunto de datos y de la cobertura de etiquetas cuando sea posible; y (iii) integración posterior del módulo MIL para predicción a nivel de lámina y selección de regiones mediante atención, con el objetivo de avanzar hacia un esquema jerárquico completo (lámina→región→célula) manteniendo trazabilidad e interpretabilidad [9, 11]. En fases posteriores se abordarán validaciones externas y aspectos de integración clínica conforme al marco regulatorio aplicable [11].

5. CONCLUSIONES

CytoLens propone un marco jerárquico para citología cervical digital orientado a triaje y priorización, con trazabilidad visual y un pipeline reproducible sobre WSI. En el estado actual, el trabajo se centra en detección de instancias basada en YOLO como mecanismo de evidencia localizada para revisión experta. Como extensión planificada del marco, se considera la incorporación de MIL para estimación a nivel de lámina y priorización de regiones mediante atención, avanzando hacia un esquema completo (lámina→región→célula). Este manuscrito se limita a describir el marco metodológico y el protocolo reproducible; la evaluación cuantitativa final, la validación externa y los estudios de integración clínica se presentarán en trabajos posteriores.

6. REFERENCIAS

1. Wu, T., Lucas, E., Zhao, F., Basu, P., & Qiao, Y. (2024). Artificial intelligence strengthens cervical cancer screening-present and future. *Cancer Biology & Medicine*, 21(10), 864-879. <https://doi.org/10.20892/j.issn.2095-3941.2024.0198>
2. Vargas-Cardona, F., Hidalgo-Tobón, S., & Tenjo-Mejía, M. (2024). Artificial intelligence for cervical cancer screening: Scoping review, 2009-2022. *International Journal of Gynecology & Obstetrics*, 165(2), 566-578. <https://doi.org/10.1002/ijgo.15179>
3. Jiang, P., Li, X., Shen, H., Chen, Y., Wang, L., Chen, H., & Gorria, P. (2023). A systematic review of deep learning-based cervical cytology screening: From cell identification to whole slide image analysis. *Artificial Intelligence Review*, 56, 2687-2758. <https://doi.org/10.1007/s10462-023-10588-z>
4. Litjens, G., Kooi, T., Ehteshami Bejnordi, B., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88. <https://doi.org/10.1016/j.media.2017.07.005>
5. Kanavati, F., Hirose, N., Ishii, T., Fukuda, A., Ichihara, S., & Tsuneki, M. (2022). A deep learning model for cervical cancer screening on liquid-based cytology specimens in whole slide images. *Cancers*, 14(5), Artículo 1159. <https://doi.org/10.3390/cancers14051159>
6. Nayar, R., & Wilbur, D. C. (2017). The Bethesda System for Reporting Cervical Cytology: A historical perspective. *Acta Cytologica*, 61(4-5), 359-372. <https://doi.org/10.1159/000475693>
7. Pangarkar, M. A. (2022). The Bethesda System for reporting cervical cytology. *CytoJournal*, 19, Artículo 28. https://doi.org/10.25259/CMAS_03_07_2021
8. Chen, H., Xu, L., & Zhou, J. (2022). Early-stage cervical cancer cell detection from cervical images using YOLOv5. *Biomedical Signal Processing and Control*, 75, Artículo 103572. <https://doi.org/10.1016/j.bspc.2022.103572>
9. Carbonneau, M.-A., Cheplygina, V., Granger, E., & Gagnon, G. (2018). Multiple instance learning: A survey of problem characteristics and applications. *Pattern Recognition*, 77, 329-353. <https://doi.org/10.1016/j.patcog.2017.10.009>
10. Wang, J., Yu, Y., Tan, Y., Wan, H., Zheng, N., He, Z., & Liu, S. (2024). Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. *Nature Communications*, 15, Artículo 4369. <https://doi.org/10.1038/s41467-024-48705-3>
11. Food and Drug Administration. (2021). Good machine learning practice for medical device development: Guiding principles. *U.S. Food and Drug Administration*. <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>

DATAMARKET: APLICACIÓN DE ANALÍTICA INTELIGENTE PARA LA OPTIMIZACIÓN Y LA FIDELIZACIÓN EN MINIMARKETS DE BARRIO

DATAMARKET: INTELLIGENT ANALYTICS APPLICATION FOR OPTIMISATION AND LOYALTY IN NEIGHBOURHOOD MINIMARKETS

Juan Manuel LÓPEZ¹, Juan Felipe ZAMBRANO², Juan CAMILO BURBANO³,
Andrés ENRÍQUEZ⁴ y Juan M. NÚÑEZ V.⁵

¹ *Universidad Autónoma de Occidente. Cali, Valle del Cauca, Colombia*

✉ juan_m.lopez_r@uao.edu.co

🌐 <https://orcid.org/0009-0005-4572-1279>

² *Universidad Autónoma de Occidente. Cali, Valle del Cauca, Colombia*

✉ juan_felipe.zambrano@uao.edu.co

🌐 <https://orcid.org/0009-0001-0810-1995>

³ *Universidad Autónoma de Occidente. Cali, Valle del Cauca, Colombia*

✉ Juan_camilo.burbano@uao.edu.co

🌐 <https://orcid.org/0009-0002-0681-215X>

⁴ *Universidad Autónoma de Occidente. Cali, Valle del Cauca, Colombia*

✉ andres_alb.enriquez@uao.edu.co

🌐 <https://orcid.org/0009-0009-8302-2072>

⁵ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jmnunez@usal.es

🌐 <https://orcid.org/0000-0002-3165-6200>

RESUMEN: Este proyecto propone una solución analítica diseñada para optimizar la operación de minimarkets de barrio mediante el uso inteligente de los datos. DataMarket aborda las limitaciones en gestión de inventarios y la falta de información sobre hábitos de compra, a partir de herramientas de segmentación de clientes, predicción de demanda y reportes automatizados. A través de inteligencia artificial y *dashboards* accesibles, DataMarket Andrés Enríquez permite mejorar la precisión en toma de decisiones, evitar el sobrestock o desabastecimiento y fortalecer la fidelización mediante promociones personalizadas. Su

implementación impulsa la transformación digital de minimarkets, lo que aumenta su rentabilidad, competitividad y sostenibilidad en el mercado.

Palabras clave: analítica de datos; gestión de inventarios; segmentación de clientes; microempresas; digitalización de procesos.

ABSTRACT: This project proposes an analytical solution designed to optimise the operation of neighbourhood mini-markets through the intelligent use of data. DataMarket addresses limitations in inventory management and the lack of information on purchasing habits, using customer segmentation tools, demand forecasting and automated reports. Through artificial intelligence and accessible dashboards, DataMarket improves decision-making accuracy, prevents overstocking or shortages, and strengthens customer loyalty through personalised promotions. Its implementation drives the digital transformation of mini-markets, increasing their profitability, competitiveness, and sustainability in the market.

Keywords: data analytics; inventory management; customer segmentation; micro-enterprises; process digitisation.

1. INTRODUCCIÓN

Los minimarkets de barrio desempeñan un papel esencial en la economía urbana, al proveer bienes de consumo cotidiano a poblaciones que dependen de la proximidad y la accesibilidad del comercio local. No obstante, estos establecimientos suelen presentar limitaciones estructurales asociadas a la ausencia de herramientas tecnológicas que respalden la toma de decisiones basada en evidencia. La gestión del inventario, la anticipación de la demanda, el entendimiento del comportamiento del cliente y la planificación comercial se realizan, mayoritariamente, mediante intuición o experiencia empírica de la persona administradora. Como consecuencia, emergen problemáticas como quiebres de stock, sobreinventario, baja rotación de productos y escasa fidelización de clientes.

En contraste, el sector retail formal ha adoptado tecnologías avanzadas de inteligencia artificial (IA), modelos predictivos y analítica de datos, lo que ha permitido incrementar la eficiencia operativa, optimizar inventarios y personalizar la experiencia del consumidor. Sin embargo, estas capacidades no han permeado a los minimarkets de barrio, debido a barreras económicas, escasa capacitación digital y falta de infraestructura tecnológica. Esta brecha digital evidencia la necesidad de soluciones accesibles que democratizen el uso de la analítica avanzada para pequeños comercios.

En este contexto se desarrolla *DataMarket*, una plataforma inteligente diseñada para proporcionar a minimarkets una herramienta accesible, intuitiva y basada en IA, capaz de transformar datos transaccionales en información accionable. La plataforma integra analítica descriptiva, segmentación automática de clientes, predicción ligera de demanda y un motor de recomendaciones basado en modelos de IA generativa y analítica estadística, adaptado a las características reales de los pequeños comercios.

El objetivo de este artículo es presentar la propuesta metodológica, conceptual y tecnológica de DataMarket, destacando su aporte innovador: *facilitar la adopción de analítica avanzada e inteligencia artificial en minimarkets urbanos mediante una solución de bajo costo, escalable y de uso sencillo*. Este trabajo expone el estado del arte, la metodología utilizada, los modelos de IA implementados, la arquitectura de la plataforma, el caso de estudio con dos empresas simuladas, los resultados esperados y las proyecciones para futuras investigaciones.

El alcance del artículo incluye diseño conceptual, experimentación con datos sintéticos, integración de IA y desarrollo funcional en plataforma web; se excluyen validaciones a gran escala con datos reales, análisis de rendimiento en ambientes productivos y comparaciones experimentales entre modelos, temas contemplados para fases posteriores del proyecto.

2. ESTADO DEL ARTE

2.1. PREDICCIÓN DE DEMANDA EN RETAIL

La literatura reciente en analítica aplicada al comercio minorista destaca la relevancia de la predicción de demanda como componente crítico para la optimización de inventarios. Modelos como ARIMA, SARIMA, Random Forest, Gradient Boosting y redes neuronales recurrentes han mostrado capacidad para reducir pérdidas operacionales, anticipar fluctuaciones estacionales y prevenir desabastecimientos [2, 4]. Sin embargo, la mayoría de estos modelos requieren grandes volúmenes de datos y capacidades computacionales elevadas, lo que dificulta su adopción en pequeños establecimientos.

2.2. SEGMENTACIÓN DE CLIENTES

El análisis del comportamiento del consumidor ha evolucionado hacia enfoques basados en análisis RFM, *clustering* no supervisado y minería de patrones. Estos métodos han demostrado incrementar la fidelización, mejorar el ticket promedio y permitir estrategias de marketing personalizadas [3]. A pesar de su efectividad, estas prácticas no son accesibles para los minimarkets debido a la falta de datos estructurados y herramientas intuitivas.

2.3. DIGITALIZACIÓN DEL COMERCIO LOCAL

Estudios recientes evidencian que la digitalización de pequeños comercios favorece la competitividad, reduce la informalidad y mejora la sostenibilidad económica [5]. No obstante, las barreras tecnológicas, económicas y educativas continúan limitando su adopción.

2.4. BRECHA TECNOLÓGICA Y OPORTUNIDAD

El análisis de la literatura muestra que, si bien existen tecnologías avanzadas aplicadas al *retail*, estas están orientadas a grandes empresas y requieren infraestructuras complejas. No existen soluciones accesibles para comerciantes de barrio que integren predicción, segmentación y analítica avanzada en un ecosistema de bajo costo.

DataMarket aborda esta necesidad ofreciendo una plataforma que adapta IA ligera, modelos interpretables y *dashboards* automatizados a la escala del minimarket.

3. METODOLOGÍA

La metodología empleada integra ingeniería de datos, diseño de sistemas y modelado de inteligencia artificial.

3.1. RECOLECCIÓN Y CARACTERIZACIÓN DE DATOS

Para el desarrollo inicial se emplearon datos sintéticos, generados para dos empresas simuladas:

- ToGo Minimarket.
- Olímpica Express.

Cada empresa recibió:

- 100 productos sintéticos.
- 100 clientes sintéticos.

Los datos fueron diseñados para presentar patrones diferenciados en:

- Precios.
- Rotación.
 - Categorías dominantes.
 - Segmentos de clientes.
 - Preferencia de compra.
- Frecuencia transaccional.

Cada *dataset* se almacena asociado al campo `company_email`, lo que permite que los *dashboards* y modelos de IA operen de manera independiente.

3.2. PROCESO ETL

El proceso ETL implementado contempla:

1. *Extracción* desde carga manual o simulación automatizada.
2. *Transformación*, que incluye:
 - Limpieza de valores nulos.

- Estandarización de categorías.
- Construcción de métricas RFM.
- Agregaciones por semana y producto.

4. CARGA EN UNA BASE DE DATOS CON ESTRUCTURA MULTIEMPRESA

4.1. ARQUITECTURA DE LA PLATAFORMA

La plataforma se implementó utilizando:

- Lovable AI para automatización de desarrollo.
- React + Tailwind.
- shadcn/ui.
- Rutas dinámicas por empresa.
- Panel administrativo independiente.
- Eliminación y creación dinámica de empresas.
- Mensajería entre administrador y empresas.
- Dashboards configurados con Recharts.
- Motor IA integrado mediante *prompt engineering*.

4.2. CONSIDERACIONES ÉTICAS

- Anonimización total de los datos sintéticos.
- Cumplimiento de buenas prácticas en privacidad.
- Transparencia en las recomendaciones generadas.

5. MODELOS DE INTELIGENCIA ARTIFICIAL UTILIZADOS EN DATAMARKET

Esta sección describe formalmente los modelos de IA que integran la plataforma.

5.1. SEGMENTACIÓN DE CLIENTES MEDIANTE K-MEANS OBJETIVO

Clusterizar clientes según comportamiento de compra con el fin de identificar patrones comerciales relevantes.

Variables utilizadas:

- Recencia
 - Frecuencia.
 - Ticket promedio.
 - Categoría preferida.

Justificación técnica: K-Means se seleccionó por:

- Su eficiencia computacional.
- Su desempeño óptimo con *datasets* pequeños.
- Su interpretabilidad para personas usuarias no técnicas.
- Su rápida actualización cuando se agregan nuevos registros.

Cada empresa genera segmentos completamente distintos debido a sus *datasets* independientes.

5.2. PREDICCIÓN LIGERA DE DEMANDA MODELO UTILIZADO

- Suavizamiento exponencial simple ($\alpha = 0,6$).
- Complementado con regresión lineal por producto.

Justificación:

- Adecuado para series cortas.
- Alta interpretabilidad.
- Bajo costo computacional.
- Robusto ante variaciones semanales en minimarkets.

Implementación: El modelo genera:

- Predicción semanal por producto.
- Alertas de riesgo de desabastecimiento.
- Estimación de sobreinventario.

5.3. MOTOR DE RECOMENDACIONES BASADO EN IA HÍBRIDA

El sistema utiliza una combinación de:

Reglas analíticas inteligentes: Generadas automáticamente a partir de:

- Rotación.
- Estacionalidad.
- Predicción de demanda.
- Segmentos K-Means.
- Ticket promedio.
- Productos críticos.

IA generativa contextual: Un modelo de lenguaje integrado en Lovable crea recomendaciones expresadas en lenguaje natural:

- Sugerencias de inventario.
- Promociones por segmento.
- Estrategias de retención.
- Recomendaciones para categorías con baja rotación.

Integración: Los *prompts* se alimentan con:

- Top productos.
- Alertas.
- Segmentos dominantes.
- Tendencias semanales.

Lo que permite que las recomendaciones sean completamente distintas para cada empresa.

6. CASO DE ESTUDIO: IMPLEMENTACIÓN EN TOGO Y OLÍMPICA

Se desarrolló un caso de estudio con dos empresas representativas. Cada una cuenta con datos sintéticos completamente diferenciados, *dashboards* propios y recomendaciones personalizadas.

Funcionalidades evaluadas

- Segmentación por empresa.
- Dashboards dinámicos.
- Motor IA adaptativo.
- Sistema multiempresa con rutas dinámicas.
- Panel administrativo con eliminación de empresas.
- Mensajería entre administrador y compañías.

7. RESULTADOS ESPERADOS

7.1. HALLAZGOS PREVISTOS

- Formación de 3-5 segmentos de clientes por empresa.
- Identificación clara de categorías de mayor rotación.
- Predicciones semanales con error controlado.
- Recomendaciones diferenciadas y coherentes por empresa.
- Disminución estimada del 15 % en quiebres de stock.
- Reducción estimada del 20 % en sobreinventario.

7.2. VISUALIZACIONES

- Tendencias de ventas.
- Rotación por categoría.
- Comparación de segmentos.
- KPI por empresa.

8. CONCLUSIONES

DataMarket constituye una solución tecnológica integral orientada a la transformación digital de los minimarkets de barrio mediante herramientas de analítica e inteligencia artificial accesibles y adaptables. La plataforma combina modelos de IA ligeros, segmentación automática, predicción de demanda y recomendaciones generativas, integrados en una arquitectura web moderna, escalable y apta para múltiples empresas simultáneas.

El sistema permite mejorar la toma de decisiones, optimizar la gestión del inventario y fortalecer la relación con el cliente, al tiempo que facilita la adopción de tecnologías avanzadas en contextos tradicionalmente excluidos de la digitalización.

9. TRABAJO FUTURO

Se proyectan las siguientes mejoras:

1. Implementación de modelos de predicción basados en *deep learning* (LSTM, Transformer).
2. Integración de análisis en tiempo real mediante *streaming* de datos.
3. Ampliación del módulo de IA generativa hacia recomendaciones automatizadas por cliente.
4. Validación empírica con datos reales de minimarkets locales.
5. Integración con POS e IoT para seguimiento de inventario en estanterías.

10. REFERENCIAS

1. Chong, A. Y. L., Li, B., Ngai, E. W. T., Ch'ng, E., & Lee, F. (2017). Predicting online product sales via online reviews, sentiments, and promotion strategies. *International Journal of Operations & Production Management*, 37(11), 1486-1508. <https://doi.org/10.1108/IJOPM-03-2016-0151>
2. Fildes, R., Ma, S., & Kolassa, S. (2022). Retail forecasting: Research and practice. *International Journal of Forecasting*, 38(3), 1161-1183. <https://doi.org/10.1016/j.ijforecast.2019.06.004>
3. Gupta, S., & Kohli, A. (2021). Enterprise retail analytics: Transforming operations and improving customer experience. *Journal of Retailing*, 97(3), 512-525.
4. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
5. Venkatesh, V., Shaw, R., & Sykes, T. (2020). Digital technology adoption in small retailers: A multi-level framework. *MIS Quarterly*, 44(3), 1165-1200. <https://doi.org/10.25300/MISQ/2020/14407>

MODELIZACIÓN INTELIGENTE PARA LA GESTIÓN SOSTENIBLE DE BIOMASA FORESTAL MEDIANTE IA GEOESPACIAL

INTELLIGENT MODELLING FOR SUSTAINABLE FOREST BIOMASS MANAGEMENT USING GEOSPATIAL AI

Sarai COBAS MERINO¹, Darío LÓPEZ RAMOS², Jairo CALDERÓN MICAÑA³ y Ana-Belén GIL-GONZÁLEZ⁴

¹ AIR Institute. Salamanca, Castilla y León, España

✉ scobas@air-institute.com

 <https://orcid.org/0009-0006-8593-4444>

² AIR Institute. Salamanca, Castilla y León, España

✉ dlopezr@air-institute.com

 <https://orcid.org/0009-0007-0164-9054>

³ AIR Institute. Salamanca, Castilla y León, España

✉ jcalderon@air-institute.com

 <https://orcid.org/0009-0008-1887-8985>

⁴ Universidad de Salamanca. Salamanca, Castilla y León, España

✉ abg@usal.es

 <https://orcid.org/0000-0001-7235-6151>

RESUMEN: El proyecto DeepForest propone un enfoque basado en la integración de inteligencia artificial y Sistemas de Información Geográfica (SIG) para apoyar la gestión sostenible de los recursos forestales mediante el aprovechamiento de la biomasa. El objetivo es identificar áreas con potencial de explotación forestal garantizando la regeneración a largo plazo de los ecosistemas. El marco metodológico combina técnicas de *machine learning* para la estimación de biomasa aérea con análisis geoespaciales orientados a la caracterización de la idoneidad forestal y la evaluación del potencial energético a escala local. Se han recopilado y preprocesado datos de inventarios forestales, variables climáticas y cobertura del suelo en la región de Castilla y León. La estimación preliminar de biomasa se aborda mediante modelos de Random Forest y XGBoost, mientras que se exploran técnicas de *clustering* no supervisado para identificar municipios con potencial para la localización de infraestructuras asociadas a

la biomasa. Los resultados iniciales muestran la viabilidad del enfoque propuesto, si bien la validación exhaustiva y la ampliación del análisis constituyen trabajo en curso.

Palabras clave: biomasa; machine learning; gestión forestal sostenible.

ABSTRACT: The DeepForest project proposes an approach based on the integration of artificial intelligence and Geographic Information Systems (GIS) to support the sustainable management of forest resources through the use of biomass. The aim is to identify areas with potential for forest exploitation while ensuring the long-term regeneration of ecosystems. The methodological framework combines machine learning techniques for estimating above-ground biomass with geospatial analyses aimed at characterising forest suitability and assessing energy potential at the local level. Data on forest inventories, climate variables and land cover in the region of Castile and León have been collected and pre-processed. Preliminary biomass estimation is addressed using Random Forest and XGBoost models, while unsupervised clustering techniques are explored to identify municipalities with potential for the location of biomass-related infrastructure. Initial results show the viability of the proposed approach, although comprehensive validation and expansion of the analysis are ongoing.

Keywords: biomass; machine learning; sustainable forest management.

1. INTRODUCCIÓN

En el contexto actual de emergencia climática y de redefinición de la soberanía energética en Europa, la biomasa forestal se ha consolidado como un recurso estratégico dentro del mix de energías renovables. La biomasa para energía siguió siendo la principal fuente renovable en la UE y representó alrededor del 59 % del consumo de energía renovable en 2021, según el último informe sobre bioenergía de la Comisión Europea [1].

La optimización del aprovechamiento de la biomasa forestal constituye, sin embargo, un desafío complejo que trasciende lo técnico y abarca dimensiones territoriales y socioeconómicas. En regiones como Castilla y León coexisten amplias áreas rurales afectadas por despoblación con zonas sometidas a elevada presión extractiva, lo que genera tensiones en la gestión sostenible de los recursos forestales. Datos oficiales muestran que el volumen de madera existente en los montes de Castilla y León alcanza cifras del orden de decenas de millones de metros cúbicos, lo que evidencia un alto potencial, pero también la necesidad de planificación y trazabilidad a escala regional [2].

A nivel normativo y estratégico, la actualización del Plan Nacional Integrado de Energía y Clima (PNIEC 2023-2030) y la Directiva de Energías Renovables (RED III) refuerzan la necesidad de digitalización, trazabilidad y planificación avanzada en el sector de la bioenergía, exigiendo criterios de sostenibilidad y verificación a lo largo de la cadena de suministro [3, 4].

Adicionalmente, la literatura y las evaluaciones europeas muestran que existe un margen amplio de disponibilidad de recurso en Europa susceptible de aprovechamiento

sostenible —con estimaciones que sitúan la fracción de bosque potencialmente disponible para suministro en porcentajes elevados a escala continental— lo que sugiere que las limitaciones operativas (fragmentación de la propiedad, falta de datos, logística) son el principal cuello de botella para incrementar tasas de aprovechamiento [5].

En este contexto, el proyecto DeepForest propone un marco metodológico basado en la integración de técnicas de inteligencia artificial, teledetección y Sistemas de Información Geográfica (SIG) para modernizar la planificación del aprovechamiento de la biomasa forestal. El objetivo es estimar biomasa aérea, identificar áreas con potencial de extracción sostenible, caracterizar la demanda energética local y optimizar la logística de abastecimiento, proporcionando herramientas de apoyo a la toma de decisiones para gestores y responsables públicos.

2. METODOLOGÍA

2.1. DATOS Y PREPROCESAMIENTO

El análisis se realizó sobre un conjunto heterogéneo de datos de Castilla y León, que incluye inventarios forestales, cobertura del suelo derivada de teledetección, variables climáticas (aridez, temperatura media, pluviometría) y registros de montes con Certificación Forestal Sostenible (públicos y privados). Todos los datos fueron normalizados y reproyectados a una malla espacial común para su integración en modelos de inteligencia artificial y análisis geoespacial.

2.2. DESCRIPCIÓN GENERAL DEL MÉTODO

El enfoque metodológico combina tres etapas principales: (i) estimación de la densidad forestal mediante segmentación de imágenes satelitales, (ii) agrupamiento de municipios según características ambientales y de cobertura forestal y (iii) cálculo de un índice de conveniencia que permite priorizar montes para un aprovechamiento sostenible de la biomasa. Las salidas incluyen un archivo .csv con atributos de los montes y un mapa interactivo que facilita la visualización y la consulta por parte de gestores territoriales.

2.3. MODELO DE DENSIDAD FORESTAL

La densidad forestal se estima mediante segmentación semántica de imágenes satelitales, formulada como predicción «pixel-wise» de la cobertura arbórea. Se evaluaron dos arquitecturas: U-Net y DeepLabV3+ (backbone ResNet-50) sobre un dataset de 800 pares imagen-máscara. Aunque DeepLabV3+ reprodujo mejor las máscaras de referencia, U-Net mostró mayor correlación con la cobertura real observada, identificando con precisión zonas arboladas y desarboladas, por lo que se seleccionó como modelo principal.

La evaluación se realizó mediante métricas de intersección sobre la unión (IoU) y F1 a nivel pixel y de parcela.

2.4. CLUSTERING Y CÁLCULO DE ÍNDICE DE CONVENIENCIA

Se aplicó *clustering* sobre vectores municipales que combinan variables climáticas y agregados de densidad forestal, descartando municipios no forestales (agricultura, urbano, cuerpos de agua). Posteriormente, se calculó un índice de conveniencia que integra variables relevantes como tipo de masa, tolerancia ecológica, pluviometría, temperatura y periodo frío. La formulación adoptada es: donde son las variables normalizadas y pesos asignables mediante técnicas de optimización. En la interpretación adoptada, valores cercanos a 0 indican mayor idoneidad para frondosas, valores cercanos a 1 para coníferas y valores intermedios para masas mixtas. Este índice se computa para montes públicos y privados certificados y sirve para filtrar y mapear franjas de prioridad de aprovechamiento sostenible.

2.5. RESULTADOS

Los resultados incluyen la selección de montes públicos y privados certificados, con sus atributos y puntaje de conveniencia, exportados a un fichero .csv. Además, se generó un mapa interactivo que muestra densidad forestal, clústeres municipales e índice de conveniencia, proporcionando una herramienta práctica de apoyo a la toma de decisiones en planificación territorial y energética.

3. PERSPECTIVAS Y PRÓXIMOS PASOS

Aunque los resultados presentados representan un avance inicial, el proyecto DeepForest seguirá desarrollando la evaluación de la demanda energética local y la estimación de la capacidad de producción de biomasa. Para ello se prevé utilizar modelos predictivos basados en Random Forest y XGBoost, junto con técnicas de optimización de rutas para mejorar la eficiencia en el abastecimiento de biomasa. La validación de las predicciones se realizará mediante datos de campo y teledetección multiespectral y LiDAR, contribuyendo a la consolidación de herramientas de planificación territorial y energética más precisas y sostenibles.

4. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y

Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

5. REFERENCIAS

1. European Commission. (2022). *Biomass for energy in the European Union (European Commission Report)*. European Commission.
2. Ministerio para la Transición Ecológica y el Reto Demográfico. (2022). *Inventario forestal de Castilla y León*. Gobierno de España.
3. Ministerio para la Transición Ecológica y el Reto Demográfico. (2023). *Plan Nacional Integrado de Energía y Clima (PNIEC) 2023-2030*. Gobierno de España.
4. European Commission. (2023). *Renewable Energy Directive (RED III)* (European Commission Report). European Commission.
5. Verkerk, P. H., Lindner, M., & Schelhaas, M.-J. (2012). Assessing forest availability for wood supply in Europe. *Forest Policy and Economics*, 17, 1-12. <https://doi.org/10.1016/j.forpol.2011.11.003>

LA INTELIGENCIA ARTIFICIAL (IA) EN LA ASISTENCIA A PERSONAS MAYORES EN ENTORNOS RURALES: REVISIÓN DEL ESTADO DEL ARTE

ARTIFICIAL INTELLIGENCE (AI) IN ASSISTING THE ELDERLY IN RURAL SETTINGS: A REVIEW OF THE STATE OF THE ART

Laura GRANDE-PÉREZ¹, Johanna MORENO-CASTELLANOS² y Ana Belén GIL-GONZÁLEZ³

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ lauragrande@usal.es

 <https://orcid.org/0009-0004-2994-8628>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jmorenoc@usal.es

 <https://orcid.org/0009-0008-3309-9891>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ abg@usal.es

 <https://orcid.org/0000-0001-7235-6151>

RESUMEN: En España, el envejecimiento poblacional y la despoblación del 80 % del territorio amenazan la sostenibilidad sanitaria. En este contexto y ante la falta de infraestructura, la inteligencia artificial (IA) es clave para mejorar la atención sanitaria en áreas rurales. Esta revisión sigue el protocolo PRISMA para analizar soluciones de IA para el cuidado de mayores.

Palabras clave: inteligencia artificial; reto demográfico.

ABSTRACT: In Spain, population ageing and depopulation of 80 % of the territory threaten healthcare sustainability. In this context and given the lack of infrastructure, Artificial Intelligence (AI) is key to improving healthcare in rural areas. This review follows the PRISMA protocol to analyse AI solutions for elderly care.

Keywords: artificial intelligence; demographic challenge.

1. INTRODUCCIÓN

La transición demográfica hacia una mayor longevidad es global. En España, el envejecimiento poblacional y la despoblación rural plantean retos críticos para la sostenibilidad sanitaria [1, 2]. La prevalencia de enfermedades crónicas y la escasez de profesionales médicos complican la atención en áreas remotas [3, 4, 5, 6].

La inteligencia artificial (IA) se perfila como solución estratégica para facilitar el envejecimiento en el domicilio (*ageing in place*), reducir la carga de especialistas y mitigar el aislamiento social [5, 7, 8]. Sus aplicaciones prometen transformar la movilidad, la accesibilidad y el diagnóstico remoto [5, 9, 10, 11].

Este estudio, utilizando el método PRISMA, revisa el estado del arte de las arquitecturas de IA, sus aplicaciones, barreras y líneas futuras en ética y usabilidad.

2. METODOLOGÍA DE LA REVISIÓN

Siguiendo el protocolo PRISMA, se consultó la base de datos *Scopus* con esta ecuación:

{TITLE-ABS-KEY («artificial intelligence» OR «AI») AND TITLE-ABS-KEY («older adults» OR «elderly» OR «personas mayores» OR «ancianos») AND TITLE-ABS-KEY («rural» OR «entornos rurales» OR «rural area»)}.

De los 107 resultados iniciales, tras aplicar filtros por tipo de documento (artículos), idioma (inglés) y relevancia temática se obtuvo un corpus de 56 estudios. Finalmente, se seleccionaron los 20 artículos más relevantes sobre la IA aplicada a la atención de personas mayores en entornos rurales.

3. APLICACIONES IA EN LA ASISTENCIA A PERSONAS MAYORES

Se ha detectado que las soluciones de IA se articulan alrededor en diferentes áreas y se describen a continuación.

Monitorización y teleasistencia: Las tecnologías de hogar inteligente e IoT, potenciadas por IA, facilitan el *ageing in place* en entornos rurales [3]. La monitorización y la detección de anomalías con sensores inalámbricos (WSN), con funcionamiento *Zero-burden*, ayudan a preservar la independencia [3]. Para superar la conectividad limitada, se emplean arquitecturas de Machine Learning y Edge AI que optimizan el procesamiento en tiempo real [4, 7], mientras que la telemedicina y la integración del 5G permiten la autogestión de enfermedades crónicas [13] y el diagnóstico remoto, supliendo la falta de especialistas [5].

Prevención y detección temprana de eventos de salud: La IA se ha demostrado efectiva en la identificación de patrones para prevenir riesgos. Algoritmos como *K-means* y *Self-Organizing Maps* (SOM) detectan cambios inusuales de actividad [12], mientras que los dispositivos

wearables previenen riesgos [14]. En salud cognitiva, tecnologías *AgeTech* y Realidad Virtual (RV) se utilizan para entrenamientos personalizados y rehabilitación [7, 13, 15]. Los sistemas de apoyo a la decisión clínica basados en Machine Learning (ML) y Deep Learning (DL) personalizan el cuidado [7].

Diagnóstico asistido y visión artificial: El DL potencia el análisis de imágenes médicas [5, 6, 11]. Aunque el análisis de imágenes dinámicas (ecografías) requiere modelos complejos, existen éxitos como la plataforma Swift Skin and Wound, una herramienta que evalúa heridas con un 99 % de consistencia, vital en comunidades rurales sin especialistas [5, 10].

Asistentes virtuales y robots sociales: Los agentes virtuales cognitivos médicos (MCVA) ofrecen recordatorios y apoyo emocional mediante procesamiento de lenguaje natural [13]. *Chatbots* basados en aprendizaje por refuerzo han demostrado mejorar la función cognitiva [15]. Complementariamente, los robots sociales brindan interacción física y emocional, aliviando la soledad [16].

Robótica de asistencia: La robótica de asistencia incluye el uso de drones equipados con brazos manipuladores para asistir en tareas cotidianas, como transportar medicamentos o suministros esenciales [17]. Esta tecnología es relevante en entornos rurales remotos donde la limitación geográfica dificulta la asistencia presencial [9].

4. DESAFÍOS ESPECÍFICOS DEL ENTORNO RURAL

La implementación de soluciones de IA en la asistencia a personas mayores en zonas rurales enfrenta desafíos que trascienden lo puramente tecnológico, siendo la brecha digital una barrera persistente [14] que, unida a la baja alfabetización digital [18] y la complejidad de los dispositivos, limita significativamente su adopción [3, 10, 14, 19]. A nivel estructural, la inestabilidad de red en las zonas rurales y el coste elevado limitan el funcionamiento de los sistemas [3, 12, 13]. En situaciones de emergencia, la insuficiencia de infraestructura digital agrava riesgos para poblaciones vulnerables [9, 20]. La aceptación por parte de los adultos mayores depende de su confianza [3]. El exceso de alarmas y la dificultad operativa pueden generar rechazo haciendo indispensable que el profesional clínico mantenga el liderazgo en la decisión final para garantizar la seguridad y la ética [21].

5. CONCLUSIONES

El uso de la IA, junto con otras tecnologías, ofrece soluciones viables para la monitorización remota (*zero-burden monitoring*), el entrenamiento cognitivo personalizado y el diagnóstico remoto especializado [5]. Sin embargo, la implementación en entornos rurales está limitada por desafíos importantes: la persistente brecha digital [11], las limitaciones de infraestructura y asequibilidad [20] y las barreras *éticas* relacionadas con la privacidad y la autonomía [9]. El *éxito* futuro depende de la innovación técnica junto con la adaptación

local y la inclusión de los usuarios en el proceso de diseño, garantizando que la tecnología refuerce la confianza y la calidad de vida en el hogar [14].

6. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

7. REFERENCIAS

1. Instituto Nacional de Estadística. (2024). *Nota de prensa: Proyecciones de población. Años 2024-2074*. INE.
2. Ministerio para la Transición Ecológica y el Reto Demográfico. (2019). *Directrices generales de la estrategia nacional frente al reto demográfico*. Gobierno de España.
3. Li, X., Zhang, C., Zhao, Y., & Feng, H. (2025). Navigating ageing in place within a smart home: Bridging stakeholder viewpoints. *Journal of Population Ageing*. Publicación anticipada en línea. <https://doi.org/10.1007/s12062-025-09432-x>
4. Lekapol, A., Raksanam, B., Sornkaew, W., Thirarattanasunthon, P., Nak-ai, W., Bunmalert, A., Kraikaew, R., Chuaykaew, B., Munira, L., & Mallongi, A. (2025). Community-driven development of the Chalerm App: A mobile health application for chronic disease management and elderly care in southern Thailand. *Journal of Multidisciplinary Healthcare*, 18, 6115-6125. <https://doi.org/10.2147/JMDH.S541792>
5. Karako, K., Song, P., Chen, Y., & Tang, W. (2020). Realizing 5G- and AI-based doctor-to-doctor remote diagnosis: Opportunities, challenges, and prospects. *BioScience Trends*, 14(5), 314-317. <https://doi.org/10.5582/bst.2020.03364>
6. Bunnell, A., Valdez, D., Wolfgruber, T. K., Quon, B., Hung, K., Hernandez, B. Y., Seto, T. B., Killeen, J., Miyoshi, M., Sadowski, P., & Shepherd, J. A. (2025). Prediction of mammographic breast density based on clinical breast ultrasound images using deep learning: A retrospective analysis. *The Lancet Regional Health - Americas*, 46, Artículo 101096. <https://doi.org/10.1016/j.lana.2025.101096>
7. Abadir, P., Oh, E., Chellappa, R., Choudhry, N., Demiris, G., Ganesan, D., Karlawish, J., Marlin, B., Li, R., Dehak, N., Arbaje, A., Unberath, M., Cudjoe, T., Chute, C., Moore, J. H., Phan, P., Samus, Q., Schoenborn, N. L., Battle, A., & Walston, J. D. (2024). Artificial intelligence and technology collaboratories: Innovating aging research and Alzheimer's care. *Alzheimer's & Dementia*, 20(4), 3074-3079. <https://doi.org/10.1002/alz.13710>
8. Keeler, L. W., & Bernstein, M. J. (2021). The future of aging in smart environments: Four scenarios of the United States in 2050. *Futures*, 133, Artículo 102830. <https://doi.org/10.1016/j.futures.2021.102830>

9. Chinbat, T., Fumihiko, N., Mihoko, M., Shinji, T., & Ryo, A. (2023). Impact assessment study of mobility-as-a-service (MaaS) on social equity through nonwork accessibility in rural Japan. *Asian Transport Studies*, 9, Artículo 100109. <https://doi.org/10.1016/j.eastsj.2023.100109>
10. Freeman, S., Sargent, M. J., Galarza, L. R., McAloney, R., & Rossnagel, E. (2025). Use of artificial intelligence-driven wound care management to enhance access to care rural and northern communities. *International Wound Journal*, 22(10), Artículo e70767. <https://doi.org/10.1111/iwj.70767>
11. Miwa, T., Minoda, R., Yamaguchi, T., Kita, S.-I., Osaka, K., Takeda, H., Kanemaru, S.-I., & Omori, K. (2022). Application of artificial intelligence using a convolutional neural network for detecting cholesteatoma in endoscopic enhanced images. *Auris Nasus Larynx*, 49(1), 11-17. <https://doi.org/10.1016/j.anl.2021.03.018>
12. Yang, D., Wang, J., Huang, L., & Wang, T. (2024). Wireless sensor network-based health monitoring and smart home assistance for the older adults in sports and wellness towns. *Frontiers in Public Health*, 12, Artículo 1399648. <https://doi.org/10.3389/fpubh.2024.1399648>
13. Mishra, S., Chaudhury, P., Tripathy, H. K., Sahoo, K. S., Jhanjhi, N., Elnour, A. A. H., & Abdelmaboud, A. (2024). Enhancing health care through medical cognitive virtual agents. *DIGITAL HEALTH*, 10, 1-15. <https://doi.org/10.1177/20552076241256732>
14. Kraft, S. A., Chopra, S., Duran, M. C., Rojina, J. A., Beretta, A., Lopez, K. I., Javan, R., Wilfond, S., Rosenfeld, M., Fogarty, J., & Ko, L. K. (2025). Perspectives of Hispanic and Latinx community members on AI-enabled mHealth tools: Qualitative focus group study. *Journal of Medical Internet Research*, 27, Artículo e59817. <https://doi.org/10.2196/59817>
15. Gao, S., & Tang, L. (2022). Research on the application of virtual reality technology in the elderly sports industry in the era of big data. *International Journal of Intelligent Systems and Applications in Engineering*, 10(4), 283-291.
16. Yun, B. H., Kim, W., Ko, H. J., Park, S. Y., Kim, J. S., Kim, G. H., Kim, B., Lee, J. H., & Kim, W. (2024). Development and effectiveness of an AI chatbot-based mobile cognitive screening and customized training application for preventing dementia: Older adults living in rural areas of South Korea. *Archives of Design Research*, 37(5), 77-90. <https://doi.org/10.15187/adr.2024.11.37.5.77>
17. Balasingam, M. (2017). Drones in medicine—The rise of the machines. *International Journal of Clinical Practice*, 71(9), Artículo e12989. <https://doi.org/10.1111/ijcp.12989>
18. Park, Y. R., & Youm, Y. (2024). Digital literacy and health of older people in a Korean rural village. *Gerontechnology*, 23(s), 1. <https://doi.org/10.4017/gt.2024.23.s.1157.opp>
19. Charlson, J. A., McGinley, E. L., Nattinger, A. B., Neuner, J. M., & Pezzin, L. E. (2017). Geographic variation of adjuvant breast cancer therapy initiation in the United States: Lessons from Medicare Part D. *Journal of the National Comprehensive Cancer Network*, 15(12), 1509-1517. <https://doi.org/10.6004/jnccn.2017.7021>
20. Kanbara, S., Shaw, R., Eguchi, K., & Das, S. (2025). Lessons from the 2024 Noto Peninsula Earthquake: Need for digital transformation in disaster response. *Progress in Disaster Science*, 25, Artículo 100400. <https://doi.org/10.1016/j.pdisas.2024.100400>
21. Goirand, M., Austin, E., & Clay-Williams, R. (2023). Mapping the flow of knowledge as guidance for ethics implementation in medical AI: A qualitative study. *PLOS ONE*, 18(11), Artículo e0288448. <https://doi.org/10.1371/journal.pone.0288448>

FOUNDATIONS AND DESIGN OF AN AI-BASED SERVICE FOR FINANCIAL AND PHARMACOLOGICAL WELL-BEING

FUNDAMENTOS Y DISEÑO DE UN SERVICIO BASADO EN INTELIGENCIA ARTIFICIAL PARA EL BIENESTAR FINANCIERO Y FARMACOLÓGICO

José ORTEGA-MOHEDANO¹ and Jaime SUÁREZ PÉREZ²

¹ *Department of Business Administration and Economics, University of Salamanca.
Salamanca, 37007, Spain*

✉ lito@usal.es

🌐 <https://orcid.org/0000-0002-4765-439X>

² *Faculty of Economics and Business, University of Salamanca. Salamanca, 37007, Spain*

✉ Jaimesuarez@usal.es

🌐 <https://orcid.org/0009-0007-4284-2926>

ABSTRACT: This paper defines a framework of relationships and processes for an electronic service aimed at supporting individuals' financial and pharmacological health, primarily through artificial intelligence and blockchain technologies, in which the theoretical foundations —legislative, technological, and conceptual— that underpin the model are established. Given the complexity of the system, the analysis focuses on educational variables —technological, financial, and health-related— and their influence on both financial and pharmacological health. The model's contribution lies in providing a reference for the design of new digital assistance services based on Big Data and AI for comprehensive personal health care. To the best of current knowledge, it is the first theoretical model to integrate relationships between these diverse health dimensions.

Keywords: technological education; health education; financial education; pharmacological health; financial health; fintech; electronic banking; insurance sector; health sector; artificial intelligence; reliable artificial intelligence.

RESUMEN: Este artículo define un marco de relaciones y procesos para un servicio electrónico destinado a apoyar la salud financiera y farmacológica de las personas, principalmente a través de tecnologías de inteligencia artificial y cadena de bloques, en el que se establecen los fundamentos teóricos —legislativos, tecnológicos y conceptuales— que sustentan el modelo.

Dada la complejidad del sistema, el análisis se centra en las variables educativas —tecnológicas, financieras y relacionadas con la salud— y su influencia tanto en la salud financiera como en la farmacológica. La contribución del modelo radica en proporcionar una referencia para el diseño de nuevos servicios de asistencia digital basados en Big Data e IA para la atención integral de la salud personal. Según los conocimientos actuales, es el primer modelo teórico que integra las relaciones entre estas diversas dimensiones de la salud.

Palabras clave: educación tecnológica; educación sanitaria; educación financiera; salud farmacológica; salud financiera; tecnología financiera; banca electrónica; sector de los seguros; sector sanitario; inteligencia artificial; inteligencia artificial fiable.

1. INTRODUCTION

The convergence of financial technology (Fintech) and health technology (Healthtech) presents an unprecedented opportunity to address systemic inefficiencies that compromise individual well-being. Currently, financial and healthcare systems operate isolated, lacking data integration or joint analysis. This fragmentation deepens «financial toxicity», where the inability to afford medication significantly impacts therapeutic adherence and health outcomes [1]. Recent studies indicate that financial toxicity correlates with higher symptom burden and suboptimal adherence across illnesses [2-4]. Concurrently, deficits in financial and health literacy limit autonomous decision-making.

This paper defines an integrated theoretical framework for an electronic service supporting individuals' financial and pharmacological health. Leveraging Artificial Intelligence (AI) and blockchain, the model provides personalized outputs while complying with EU AI Act (2024/1689).

2. LEGISLATIVE AND TECHNOLOGICAL FRAMEWORK

The model rests on four pillars: legislative, technological, financial theory, and pharmacological theory.

Legislative Framework. The EU AI Act (2024/1689) classifies AI systems evaluating creditworthiness or establishing risk for life and health insurance as «high-risk» [5]. Our framework incorporates these requirements by design, ensuring traceability, transparency, and human oversight. Data handling adheres to GDPR and HIPAA standards, protecting sensitive personal data.

Technological Infrastructure. The architecture integrates AI for predictive analysis and blockchain for security and auditability. AI algorithms process heterogeneous data to predict financial stress and medication non-adherence, increasing productivity [6]. Blockchain ensures data integrity and traceability, mitigating fraud and ensuring that health records [7]. It facilitates smart contracts that can automate compliance and insurance claims [8].

Educational Foundations. Financial literacy determines well-being, influencing risk aversion and the effective use of financial instruments. Digital financial literacy significantly enhances financial well-being [9]. Similarly, health literacy is critical for treatment success; low literacy correlates with poor adherence. The service employs patient stratification to tailor interventions, as patients with high health literacy achieve better outcomes [10].

3. INTEGRATED FRAMEWORK OF FINANCIAL HEALTH AND PHARMACOLOGICAL HEALTH

The core contribution of this research is a proposal of an *Integrated Framework of Financial Health and Pharmacological Health* (see Figura 1), which details the flow from data ingestion to personalized output.

Integrated Inputs. The system aggregates data from two traditionally separate domains: (1) *Financial Inputs*, such as credit history, bank transactions, investment portfolio, and financial literacy indicators; and (2) *Healthcare Inputs*, based on electronic health records (EHR), prescriptions, adherence patterns, and demographics.

The Processing Layer. The processing core utilizes AI to identify correlations between financial capacity and health requirements. For instance, it analyzes if a patient's liquidity issues are causing medication rationing. Blockchain mechanisms ensure that this data exchange is secure and traceable, addressing the «black box» problem of AI transparency [11].

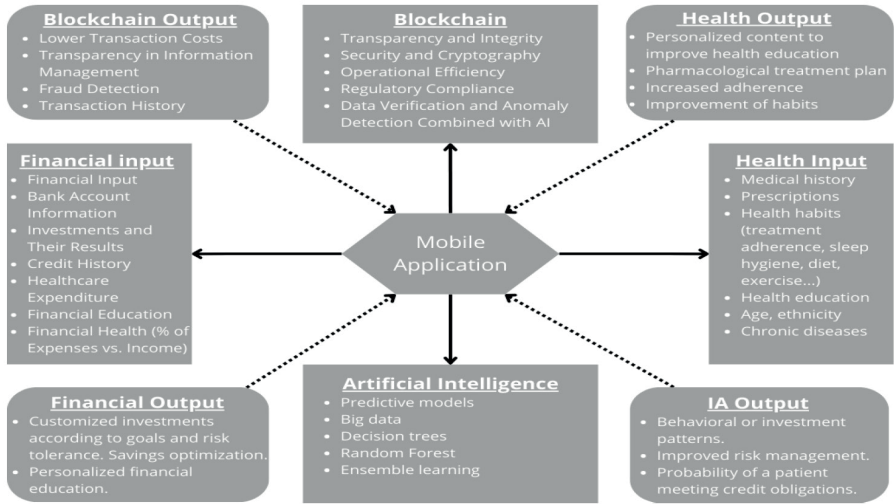


Figura 1. Integrated Conceptual Framework: The mobile application integrates Blockchain, AI, and Big Data. Inputs (Financial/Healthcare) are processed via predictive models to generate personalized outputs (Investment advice, Treatment plans), mediated by educational enhancements.

Personalized Outputs. The system generates three types of actionable, personalized recommendations: (1) *Financial*, such as risk-adjusted investment plans or savings strategies specifically designed to cover healthcare costs; (2) *Pharmacological*, like precision treatment plans that consider financial feasibility, thereby reducing abandonment rates; and (3) *Educational*, that is, targeted content to bridge literacy gaps in both domains.

4. DISCUSSION: THE FINANCIAL-HEALTH NEXUS

Comparative Analysis. While both sectors use decision trees, their objectives differ: finance maximizes assets via risk profiling, while pharmacology optimizes health via pathology segmentation. Our model demonstrates interdependence: improved financial health facilitates treatment access, while physical health enhances economic productivity [12].

Novelty and Impact. To our knowledge, no peer-reviewed literature has previously integrated financial and pharmacological health as interdependent variables managed by a single trustworthy AI service. By treating education as a transversal axis, the model empowers users rather than merely dictating actions.

5. CONCLUSIONS

This paper establishes the foundations for a Trustworthy AI electronic service bridging financial stability and pharmacological adherence. Aligned with the EU AI Act, it ensures safe, transparent high-risk AI applications. Future research should focus on prospective validation through randomized controlled trials (RCTs) to measure impact on «financial toxicity» and clinical outcomes.

6. ACKNOWLEDGEMENTS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

7. REFERENCES

1. Carrera, P. M., Kantarjian, H. M., & Blinder, V. S. (2018). The financial burden and distress of patients with cancer: Understanding and stepping-up action on the financial toxicity of cancer treatment. *CA: A Cancer Journal for Clinicians*, 68(2), 153-165. <https://doi.org/10.3322/caac.21443>
2. Rana, J., Shah, Z., Shah, S., Khan, N., & Ahmad, I. (2025). The financial toxicity of cancer: Unveiling global burden and risk factors - a systematic review and meta-analysis. *BMJ Global Health*, 10(2), Artículo e017133. <https://doi.org/10.1136/bmjgh-2024-017133>
3. Patel, R., Sharma, A., Kumar, S., & Singh, R. (2025). Financial toxicity in diabetes: The state of what we know. *Current Diabetes Reports*, 25, Artículo 32. <https://doi.org/10.1007/s11892-025-01588-0>
4. Slavin, S. D., Khera, R., Zafar, S. Y., Nasir, K., & Warraich, H. J. (2021). Financial burden, distress, and toxicity in cardiovascular disease. *American Heart Journal*, 238, 75-84. <https://doi.org/10.1016/j.ahj.2021.04.011>
5. European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. *Official Journal of the European Union*, L, Artículo 2024/1689.
6. Adnan, H., Shidani, A., Clifton, L., Bankhead, C., & Perera-Salazar, R. (2025). Implementation framework for AI deployment at scale in healthcare systems. *iScience*, 28(2), Artículo 108253. <https://doi.org/10.1016/j.isci.2025.112406>
7. Tagde, P., Tagde, S., Bhattacharya, T., Rahman, M. H., & Gorria, P. (2021). Blockchain and artificial intelligence technology in e-Health. *Environmental Science and Pollution Research*, 28, 52810-52831. <https://doi.org/10.1007/s11356-021-16223-0>
8. Tanwar, S., Parekh, K., & Evans, R. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50, Artículo 102407. <https://doi.org/10.1016/j.jisa.2019.102407>
9. Kamble, P., Mehta, A., & Rani, N. (2023). Financial inclusion and digital financial literacy: Do they matter for financial well-being? *Social Indicators Research*, 171, 777-807. <https://doi.org/10.1007/s11205-023-03264-w>
10. Chalasani, S., Syed, J., Ramesh, M., Barman, M., & Kumar, S. (2023). Artificial intelligence in the field of pharmacy practice: A literature review. *Exploratory Research in Clinical and Social Pharmacy*, 12, Artículo 100368. <https://doi.org/10.1016/j.rcsop.2023.100346>
11. Rooson, D., Tatla, V., Li, Y., Rooson, M. R., & Gorria, P. (2022). Framework to enable pharmacist access to health care data using Blockchain technology and artificial intelligence. *Journal of the American Pharmacists Association*, 62(5), 1362-1368. <https://doi.org/10.1016/j.japh.2022.02.018>
12. Panos, G.A., & Wilson, J. O. S. (2020). Financial literacy and responsible finance in the FinTech era: Capabilities and challenges. *The European Journal of Finance*, 26(4-5), 297-301. <https://doi.org/10.1080/1351847X.2020.1717569>

UNA HERRAMIENTA UNIFICADA PARA ESTUDIAR LOS CULTIVOS FORESTALES DE ESPAÑA

A UNIFIED TOOL FOR STUDYING FOREST CROPS IN SPAIN

Jaime GABRIEL VEGAS¹, Maider Araceli URBON JIMÉNEZ², Ana DE LUIS REBOREDO³,
Belén PÉREZ LANCHO⁴ y Ana-Belén GIL-GONZÁLEZ⁵

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ JaimeGabrielVegas@usal.es

 <https://orcid.org/0009-0006-8923-5405>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ murbon001@usal.es

 <https://orcid.org/0009-0002-6308-4398>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ adeluis@usal.es

 <https://orcid.org/0000-0001-5354-9054>

⁴ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ lancho@usal.es

 <https://orcid.org/0000-0002-0934-8316>

⁵ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ abg@usal.es

 <https://orcid.org/0000-0001-7235-6151>

RESUMEN: Este trabajo presenta una herramienta integral diseñada para modelar y predecir la captura de carbono en los cultivos forestales de España, con el fin de optimizar la gestión de créditos de carbono y proyectos de reforestación. A partir del procesamiento de datos del Inventario Forestal Nacional (IFN), combinados con variables climáticas e índices vegetales, se entrenaron diversos modelos de aprendizaje automático. Los resultados destacan el rendimiento del algoritmo CatBoost, alcanzando un $R^2 > 0,84$ en la estimación de toneladas de carbono. El estudio valida la metodología mediante un caso práctico y aporta como contribución fundamental una base de datos relacional estructurada que resuelve la dispersión de la información pública.

*Jaime Gabriel Vegas, Maider Araceli Urbon Jiménez, Ana de Luis Reboredo,
Belén Pérez Lancho y Ana-Belén Gil-González
Una herramienta unificada para estudiar los cultivos forestales de España*

Palabras clave: bosques españoles; captura de carbono; reforestación; machine learning; créditos de carbono; Inventario Forestal Nacional.

ABSTRACT: This paper presents a comprehensive tool designed to model and predict carbon capture in forest crops in Spain, with the aim of optimizing the management of carbon credits and reforestation projects. Based on data processing from the National Forest Inventory (IFN), combined with climate variables and vegetation indices, various machine learning models were trained. The results highlight the performance of the CatBoost algorithm, achieving an $R^2 > 0.84$ in the estimation of carbon tons. The study validates the methodology through a case study and makes a fundamental contribution in the form of a structured relational database that resolves the dispersion of public information.

Keywords: spanish forests; carbon capture; reforestation; machine learning; carbon credits; National Forest Inventory.

1. INTRODUCCIÓN

Durante las últimas décadas, los bosques españoles han experimentado transformaciones profundas y, en ciertos aspectos, alarmantes. Si bien la superficie forestal ha aumentado debido al éxodo rural y la expansión natural de la vegetación, este crecimiento no se ha traducido necesariamente en una mejora ecológica. Persisten problemas estructurales como una diversidad limitada, la homogeneización del paisaje y la predominancia de masas monoespecíficas, las cuales presentan una alta vulnerabilidad frente a eventos catastróficos, como por ejemplo incendios [1, 3].

España se sitúa entre las regiones europeas con mayor exposición al cambio climático [6], manifestándose este fenómeno en temporadas de incendios más extensas o en una reducción en la disponibilidad hídrica [4], entre otros efectos. Estas presiones comprometen la estabilidad de los ecosistemas forestales.

En este escenario, el papel de los bosques como sumideros naturales de carbono adquiere un valor estratégico [2], al constituir la captura y el almacenamiento de CO_2 en la biomasa y el suelo un pilar fundamental de las políticas climáticas europeas. Para garantizar la credibilidad de estos mecanismos financieros y ecológicos, resulta imprescindible disponer de datos precisos y modelos robustos que cuantifiquen con rigor el carbono capturado, así como modelos predictivos que permitan orientar la toma de decisiones.

2. METODOLOGÍA

El objetivo principal es el desarrollo de una herramienta integral que permita optimizar la amortización de créditos de carbono por (re)forestación. Para ello, se ha modelado la absorción de CO_2 en cultivos forestales españoles, identificando las variables críticas que influyen en el desarrollo arbóreo.

2.1. FUENTES DE DATOS

El núcleo del conjunto de datos proviene del Inventario Forestal Nacional (IFN) [5], un registro exhaustivo y de dominio público de la evolución desde 1986 del estado de los bosques españoles. Este inventario proporciona información de campo de alto valor, incluyendo:

- *Variables edáficas* que caracterizan los suelos forestales.
- *Variables biológicas*. Conteo de pies por especie y clase diamétrica y su caracterización biológica (origen, estado...).

La predicción se realiza a nivel de parcela y especie arbórea, abarcando un conjunto de datos de aproximadamente 50 000 parcelas y cerca de 200 especies. Dado que los datos del IFN se encuentran georreferenciados, ha sido posible enriquecer el modelo cruzando esta información con bases de datos meteorológicas e índices de vegetación satelitales.

3. RESULTADOS

Se llevó a cabo una comparativa exhaustiva de diversos algoritmos de aprendizaje automático, obteniendo el modelo basado en *CatBoost* el mejor rendimiento: explica más del 84 % de la variabilidad ($R^2 > 0,84$) en las toneladas de carbono capturadas por cultivo forestal.

El análisis de errores muestra que las desviaciones medias del modelo son bajas en comparación con la variabilidad natural de la variable objetivo. Esto indica que, a pesar de la alta heterogeneidad inherente a las parcelas forestales, el modelo mantiene un error promedio estable, validando su utilidad para estimaciones a largo plazo (Tabla 1).

Tabla 1. Métricas del modelo y estadísticos de la variable objetivo: carbono (aéreo y radical) en toneladas capturado por el cultivo forestal

Métricas		Estadísticos	
Indicador	Valor	Indicador	Valor
R^2	0,845	Media	24,6 tC
RMSE	13,8 tC	Desv. est.	35,8 tC
MAE	6,7 tC	Máximo	420,5 tC

3.1. CASO DE ESTUDIO: ÁLAVA

A modo de validación práctica, se presenta un caso de estudio en una parcela ubicada en la provincia de Álava (Figura 1). Se compararon plantaciones de pino y roble con una densidad de 500 pies/ha y un horizonte de predicción de 30 años.



Figura 1. Parcela de estudio ubicada en la provincia de Álava

El análisis automático de la ubicación identifica un clima húmedo, pendiente moderada (10-15°) y orientación oeste. Los indicadores satelitales corroboran una alta productividad histórica, con índices EVI y NDII favorables. Los resultados (Tabla 2) muestran mayor captura para pinos ($\approx 66 \text{ tCO}_2$) frente a robles ($\approx 54 \text{ tCO}_2$), coherente con las tasas de crecimiento de coníferas en zonas húmedas.

4. CONCLUSIONES

Los resultados confirman la viabilidad del objetivo planteado, demostrando que se dispone de recursos suficientes a nivel nacional para este tipo de modelado. Una contribución fundamental del proyecto ha sido el procesamiento del Inventario Forestal Nacional (IFN). Si bien estos datos del MITECO son públicos y de gran calidad, su formato original presenta barreras significativas (dispersión, nomenclatura confusa).

Tabla 2. Predicción comparativa de captura de CO_2 a 30 años

Especie	Densidad (pies/ha)	Captura est. (tCO_2)
Pino (<i>Pinus</i>)	500	≈ 66
Roble (<i>Quercus robur</i>)	500	≈ 54

Este trabajo ha solventado dichas dificultades mediante la recopilación, depuración, reestructuración relacional y documentación exhaustiva de la información, transformando un recurso complejo en una base de datos SQL accesible y manejable para futuros trabajos de investigación.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA). Referencia: TSI-100933-2023-0001. Financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (NextGenerationEU).

6. REFERENCIAS

1. Agencia SINC. (2016, 17 de marzo). *La homogeneización de los árboles reduce la funcionalidad de los bosques*. <https://www.agenciasinc.es/Noticias/La-homogeneizacion-de-los-arboles-reduce-la-funcionalidad-de-los-bosques>
2. European Commission. (2021). *New EU forest strategy for 2030* (COM(2021) 572 final). European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0572>
3. European Commission. (2022, 5 de octubre). *Abandoned agriculture in Valencia region (Spain): Transition to woodland and scrubland over a 50 year period*. Science for Environment Policy. https://environment.ec.europa.eu/news/abandoned-agriculture-valencia-region-spain-transition-woodland-and-scrubland-over-50-year-period-2022-10-05_en
4. Masson-Delmotte, V., & Zhai, P. (2022). *Tendencias regionales de los fenómenos extremos en el informe del IPCC de 2021*. Grupo Intergubernamental de Expertos sobre el Cambio Climático.
5. Ministerio para la Transición Ecológica y el Reto Demográfico. (2023). *Inventario forestal nacional: Memoria metodológica*. Gobierno de España. <https://www.miteco.gob.es/content/dam/mitco/es/biodiversidad/estadisticas/inventarios-nacionales/IFN-Metodolog%C3%ADa.pdf>
6. Sanz, M. J., & del Castillo, E. G. (2020). *Impactos y riesgos derivados del cambio climático en España*. Ministerio para la Transición Ecológica y el Reto Demográfico.

HEURISTIC AND EVOLUTIONARY OPTIMIZATION FOR UWB ANCHOR PLACEMENT

OPTIMIZACIÓN HEURÍSTICA Y EVOLUTIVA PARA LA COLOCACIÓN DE ANCLAJES UWB

Matei ANDREI MADALIN¹ and Lenuta ALBOAIE²

¹ *University Alexandru Ioan Cuza of Iasi, Department of Computer Science, Romania*
✉ bmatei.madalin@info.uaic.ro
 <https://orcid.org/0009-0003-5629-9857>

² *University Alexandru Ioan Cuza of Iasi, Department of Computer Science, Romania*
✉ lenuta.alboaie@info.uaic.ro
 <https://orcid.org/0000-0003-0969-8102>

ABSTRACT: Accurate and semantically-aware indoor navigation presents a significant challenge for existing assistive technologies (AT), particularly for visually impaired users. This research proposes a comprehensive framework that integrates Ultra-Wideband (UWB) for high-precision localization with LiDAR for detailed 3D environment mapping. We present a methodology for optimizing UWB anchor placement that minimizes Geometric Dilution of Precision (GDOP). We implement and compare two distinct optimization strategies for this task: a Hill-Climbing (HC) algorithm and a Genetic Algorithm (GA). Both methods seek to maximize a novel objective function derived from the 3D semantic model, which rewards broad coverage, penalizes anchor overlap, and enforces Line-of-Sight (LoS). The HC provides an iterative path to a local optimum, while the GA explores a wider solution space to identify a potentially superior, globally optimal configuration.

Keywords: UWB localization; assistive technology; indoor navigation; visually impaired people.

RESUMEN: La navegación interior precisa y sensible al contexto semántico supone un reto importante para las tecnologías de asistencia (TA) existentes, especialmente para los usuarios con discapacidad visual. Esta investigación propone un marco integral que integra la banda ultraancha (UWB) para una localización de alta precisión con LiDAR para un mapeo detallado del entorno en 3D. Presentamos una metodología para optimizar la colocación de anclajes UWB que minimiza la dilución geométrica de la precisión (GDOP). Implementamos

y comparamos dos estrategias de optimización distintas para esta tarea: un algoritmo Hill-Climbing (HC) y un algoritmo genético (GA). Ambos métodos buscan maximizar una nueva función objetivo derivada del modelo semántico 3D, que recompensa la amplia cobertura, penaliza la superposición de anclajes y aplica la línea de visión (LoS). El HC proporciona una ruta iterativa hacia un óptimo local, mientras que el GA explora un espacio de soluciones más amplio para identificar una configuración potencialmente superior y globalmente óptima.

Palabras clave: localización UWB; tecnología de asistencia; navegación en interiores; personas con discapacidad visual.

1 INTRODUCTION

Ultra-Wideband (UWB) technology has emerged as the standard for high-precision indoor localization due to its resilience to multipath interference. However, the accuracy of UWB is fundamentally constrained by the geometry of the anchor deployment. We measure this using the Geometric Dilution of Precision (GDOP) [1]. A poor geometric arrangement—such as placing anchors collinearly or in positions obstructed by furniture—can degrade sub-decimeter accuracy to meter-level errors, rendering the system unsafe for blind users.

The challenge lies in the complexity of modern indoor spaces. Manually determining the optimal 3D placement of anchors in a furnished room to ensure Line-of-Sight (LoS) and minimize GDOP is a hard problem for human installers.

This paper addresses this problem by automating the infrastructure design. We propose a pipeline that utilizes consumer LiDAR (via smartphones) to generate a semantic 3D model of the environment. We then formulate the anchor placement problem as a multi-objective optimization task. The core contribution of this work is the implementation and comparative analysis of two optimization approaches: a fast heuristic method (Hill Climbing) and a robust evolutionary method (Genetic Algorithm). We demonstrate that, while heuristics are computationally efficient, the evolutionary approach provides the global optimization needed to ensure user safety in complex geometric configurations.

2. RELATED LITERATURE

Traditional approaches often rely on simplified 2D floor plans, ignoring the verticality of obstacles which is critical for 3D UWB [2] positioning.

Heuristic methods, such as Simulated Annealing or Force-Directed algorithms, have been applied to maximize coverage in WSNs. These methods are generally fast but prone to entrapment in local optima, particularly in non-convex environments (e.g., L-shaped rooms).

Evolutionary computation, specifically Genetic Algorithms (GA) [3], has been identified as a powerful tool for discrete and continuous optimization problems where the search space is vast and rugged. Holland's schema theory suggests that GAs can

process building blocks of solutions to locate global optima. In the context of UWB, limited research has combined 3D semantic obstacles (furniture, walls) with evolutionary optimization to minimize GDOP explicitly. This work bridges that gap by applying a real-valued Genetic Algorithm directly to a semantic digital twin of the environment.

3. SYSTEM AND METHODOLOGY

3.1. PROBLEM FORMULATION (THE OBJECTIVE FUNCTION)

We define the fitness of a specific anchor configuration via a multi-objective function . The goal is to maximize :

Where:

- *Coverage (C)*: The number of sampled points in the room volume that fall within the effective range of the anchors.
- *Overlap Penalty (O)*: A penalty applied if the Euclidean distance between any two anchors is less than 1.0 m. This enforces geometric diversity.
- *Line-of-Sight Penalty (L)*: The most critical factor for UWB accuracy. We use ray-casting to check the segment between every pair of anchors. If a ray intersects a semantic obstacle (furniture/wall), a heavy penalty is applied.

3.2. OPTIMIZATION ALGORITHMS

To solve the anchor placement problem defined by the objective function (P), we initially employed a Stochastic Hill Climbing [4] algorithm. This trajectory-based heuristic proved highly effective for simple, single-room environments where the search space is relatively small and the number of anchors is low. In these constrained scenarios, the algorithm rapidly converges to a satisfactory geometric configuration by iteratively perturbing the position of a single anchor and accepting the change if it improves the overall fitness score. However, as the scope of the problem expanded to multi-room environments requiring a network of more than 20 anchors, the limitations of this greedy approach became apparent. The search landscape in such complex topologies is rugged, characterized by numerous local optima created by walls, furniture, and exclusion zones. The Hill Climbing algorithm, lacking a mechanism to accept temporary performance degradation or explore distant regions of the search space, frequently stagnated in suboptimal configurations, failing to bridge the geometric gaps between separate rooms.

To overcome the tendency of heuristic methods to become trapped in local optima within high-dimensional search spaces, we implemented a Genetic Algorithm (GA). Unlike the single-solution approach of Hill Climbing, the GA evolves a population of simultaneous candidate solutions, allowing it to explore the global search space more effectively. We adapted standard evolutionary theory to our specific domain by utilizing a real-valued representation, where each chromosome consists of a list of 3D vectors corresponding to

the physical coordinates of the anchors. This direct mapping ensures that the genetic operations remain semantically meaningful within the context of the physical room geometry. The evolutionary process is driven by tournament selection, which identifies high-performing parent configurations while maintaining sufficient population diversity to prevent premature convergence.

The algorithm employs two primary operators to drive optimization. First, a uniform crossover operator recombines the spatial features of two parent configurations, allowing the system to merge, for instance, a strong geometric arrangement in one room from the first parent with an optimal corridor placement from the second. Second, to prevent the stagnation observed in the heuristic approach, we utilized a random resetting mutation operator. This mechanism introduces stochastic diversity by randomly relocating an anchor to a new valid position within the candidate zones, effectively forcing the algorithm to jump out of local optima and explore new geometric possibilities. To ensure that the quality of the solution never regresses across generations, an elitism strategy preserves the highest-performing individuals. Our experimental analysis demonstrates that in complex scenarios involving large anchor networks, the Genetic Algorithm consistently identifies global optima that are superior to the results produced by Hill Climbing, justifying the additional computational cost with significantly improved localization accuracy and coverage. Results are presented in Figure 1.

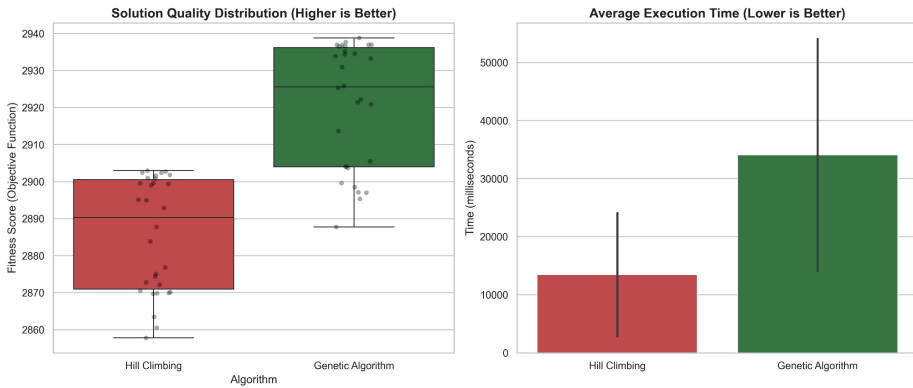


Figure 1. Performance comparison between Hill Climbing (red) and Genetic algorithm (green)

4. CONCLUSIONS AND FUTURE DEVELOPMENT

This paper presented a methodology for automating the deployment of UWB assistive navigation systems. By integrating semantic 3D mapping with evolutionary optimization, we addressed the critical dependency of UWB accuracy on anchor geometry.

Our comparative analysis demonstrates that while heuristic methods like Hill Climbing are computationally efficient, they are insufficient for complex indoor environments where safety is paramount. The Genetic Algorithm proved superior in navigating the rugged landscape of the objective function, reliably avoiding local optima to minimize GDOP.

For future development, we aim to extend this framework to multi-room environments, where the GA must optimize «hand-over» zones between different anchor clusters. Additionally, we plan to integrate «User-in-the-Loop» feedback, allowing the optimization process to account for specific user traversal patterns within the home.

5. REFERENCES

1. Yang, B., Yang, E., Yu, L., & Loeliger, A. (2022). High-precision UWB-based localisation for UAV in extremely confined environments. *IEEE Sensors Journal*, 22(1), 1020-1029. <https://doi.org/10.1109/JSEN.2021.3130724>
2. Al-Okby, M. F. R., Junginger, S., Roddelkopf, T., & Thurow, K. (2024). UWB-based real-time indoor positioning systems: A comprehensive review. *Applied Sciences*, 14(23), Artículo 11005. <https://doi.org/10.3390/app142311005>
3. Forrest, S. (1996). Genetic algorithms. *ACM Computing Surveys*, 28(1), 77-80. <https://doi.org/10.1145/234313.234350>
4. Selman, B., & Gomes, C. P. (2006). Hill-climbing search. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (1st ed.). Wiley. <https://doi.org/10.1002/0470018860.s00015>

EVALUATING THE IMAGE QUALITY OF MAMMOGRAPHY PHANTOM IMAGES

EVALUACIÓN DE LA CALIDAD DE IMAGEN EN IMÁGENES DE FANTOMA MAMOGRÁFICO

Valentina CORCHUELO^{1,3}, Juan MANUEL NÚÑEZ³, Edison SALAZAR HURTADO² and Hernán Darío BENÍTEZ RESTREPO⁴

¹ *Pontifical Javeriana University Cali, Department of Electronics and Computer Science. Cali, Colombia*

✉ valentina.corchuelo@javerianacali.edu.co

🌐 <https://orcid.org/0000-0001-7429-7634>

² *Imbanaco Clinic, Quiron Health. Cali, Colombia*

🌐 <https://orcid.org/0009-0006-5418-0213>

³ *BISITE Research Group, University of Salamanca. Salamanca, Spain*

✉ jmnunez@usal.es

🌐 <https://orcid.org/0000-0002-3165-6200>

⁴ *Edenmed, Colombia*

🌐 <https://orcid.org/0000-0002-2661-8867>

ABSTRACT: Image quality assessment in medical imaging constitutes a relevant contribution to improving the reliability and productivity of healthcare processes. In this work, an approach based on the acquisition of mammography phantom images is presented to evaluate their quality and analyze the relationship with image quality assessment tasks performed by human observers. A dataset was constructed from X-ray images acquired on a mammography system using an accreditation mammography phantom under different exposure conditions. On these images, objective image quality features such as signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were analyzed. Additionally, a tool for subjective image evaluation was developed, in which expert observers rated the visibility of the phantom structures. The proposed approach aims to support the development of objective image quality measures that are consistent with the performance of visual structure detection in quality control images for mammography equipment.

Keywords: image quality assessment; accreditation phantom; mammography; SNR/CNR.

RESUMEN: La evaluación de la calidad de la imagen en el ámbito de la imagen médica constituye una contribución relevante para mejorar la fiabilidad y la productividad de los procesos sanitarios. En este trabajo se presenta un enfoque basado en la adquisición de imágenes fantasma de mamografías para evaluar su calidad y analizar la relación con las tareas de evaluación de la calidad de la imagen realizadas por observadores humanos. Se construyó un conjunto de datos a partir de imágenes de rayos X adquiridas en un sistema de mamografía utilizando un fantoma de mamografía acreditado en diferentes condiciones de exposición. En estas imágenes se analizaron características objetivas de calidad de imagen, como la relación señal-ruido (SNR) y la relación contraste-ruido (CNR). Además, se desarrolló una herramienta para la evaluación subjetiva de imágenes, en la que observadores expertos calificaron la visibilidad de las estructuras del fantoma. El enfoque propuesto tiene como objetivo apoyar el desarrollo de medidas objetivas de calidad de imagen que sean coherentes con el rendimiento de la detección visual de estructuras en imágenes de control de calidad para equipos de mamografía.

Palabras clave: Evaluación de la calidad de imagen; fantoma de acreditación; mamografía; SNR/CNR.

1. INTRODUCTION

Breast cancer represents a major public health problem in Colombia, with approximately 2,226 women dying each year from this disease [1]. According to [2] in Colombia, early detection of breast cancer is based on three key activities: undergoing a mammogram (MG) every two years in women between 50 and 69 years of age, or earlier at the physician's discretion; conducting an annual clinical breast examination in women aged 40 years and older as part of routine physical assessments; and promoting monthly breast self-examination as a self-care practice.

In Colombia, the Ministry of Health has adopted several initiatives aimed at reducing the incidence, mortality, and disability associated with breast cancer, including the promotion of timely, accurate, and coordinated interventions within the healthcare system [3, 4]. Among the strategies currently promoted by the government is monitoring the quality of mammography services across the country.

To perform quality control procedures on mammography equipment, a phantom is used, consisting of a block of polymethyl methacrylate (PMMA) that simulates the X-ray attenuation of an average compressed breast and contains embedded structures. System performance is evaluated by radiology teams through human visual identification of visible objects for each structure type, including fibers, microcalcifications, and masses. Figura 1 shows a reference image of the ACR RMI 156.

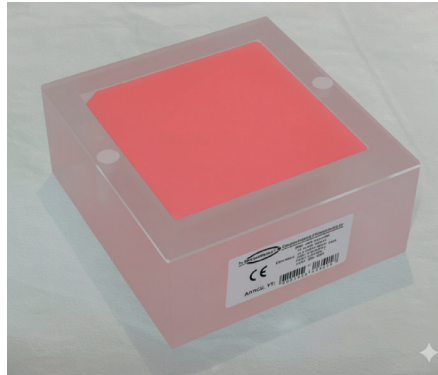


Figura 1. Phantom ACR RMI 156 [5]

In this work, through the integration of the dataset and expert opinions, we aim to establish a basis for the development of automated solutions that support quality control procedures for mammography equipment and contribute to the early detection of breast cancer.

2. MATERIALS AND METHODS

The research was conducted under a quantitative, descriptive, and correlational approach. In the first stage, a dataset of mammography phantom (MG) images was prospectively constructed at Clínica Imbanaco, a high-complexity healthcare institution located in Cali, Colombia. Subsequently, the structures present in the images were analyzed and a subjective study was conducted, in which a group of evaluators consisting of medical imaging technologists, medical physicists, radiologists, and biomedical and electronic engineers with experience in medical imaging, from different institutions in the city of Cali, performed structure scoring and perceptual quality assessment on an assigned dataset.

To conduct the subjective study, the project was approved by the Health Research Ethics Committee of the Faculty of Health Sciences at Pontificia Universidad Javeriana Cali.

3. ANALYSIS AND RESULTS

A custom graphical user interface (GUI) was developed in MATLAB®, consisting of four modules. The first module allowed the registration of the evaluator's identifier for the study; the second and third modules displayed images for each evaluator to identify the number of disks, fibers, and microcalcification groups; and the fourth module enabled participants to rate the perceptual quality of each mammography phantom image (PMGI).

Previous work by the research group analyzed objective image quality metrics in mammography phantom images (PMGI), such as signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), and natural scene statistics (NSS), under different acquisition conditions, establishing an initial quantitative characterization of image quality [6]. The subjective study included a dataset composed of ten undistorted PMGIs with high CNR values, thirty distorted PMGIs, five undistorted projected PMGIs, and thirty-five distorted and projected PMGIs. Here, the projections refer to spatial transformations of the PMGIs applied to alter the apparent location of structures. Projections were performed in vertical (V), horizontal (H), and diagonal (D) directions in order to minimize bias derived from the spatial memory of object locations. The complete dataset consisted of eighty PMGIs. Figura 2 shows the flow diagram of the tool.

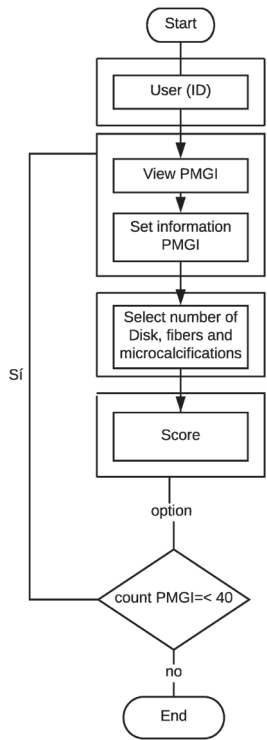


Figura 2. Tool diagram

The evaluations were conducted by twenty-one experts in two 30-minute sessions under controlled lighting conditions, using a Spyder5 PRO device to calibrate the display according to industry color standards. All expert evaluations were recorded through the GUI for subsequent data processing and analysis.

4. CONCLUSIONS

Since the CNR of PMGIs is related to the detection of structures in the image, and the attenuation coefficient of microcalcifications is higher than that of fiber and disk-type structures, this factor was fundamental in the construction of the dataset for the subjective study.

The evaluation of PMGIs is associated with the selection of disk, fiber, and microcalcification group structures. According to the time required to perform the task, it was identified that evaluators may need between 20 and 40 minutes to assess 40 images. Due to the repetitive selection of structures in PMGIs, visual fatigue may also occur among evaluators.

The MATLAB® based graphical user interface (GUI) enabled the integration of structure detection information from PMGIs and the assignment of a quality score to each image. These results were consolidated to support subsequent data processing stages, such as the analysis and modeling of image quality with the aim of facilitating the prediction of diagnostic image quality.

5. ACKNOWLEDGMENTS

To Pontificia Universidad Javeriana Cali, Clínica Imbanaco, to health care professionals of different health care centers who participated in the subjective study and cosupervised the project.

6. REFERENCES

- Ramos, C. P., & Duarte, R. C. (2015). *Incidencia, mortalidad y prevalencia de cáncer en Colombia, 2007-2011*. Ministerio de Salud y Protección Social.
- Ministerio de Salud y Protección Social. (2017). *Guía de práctica clínica para la detección temprana, tratamiento integral, seguimiento y rehabilitación del cáncer de mama* (2.^a ed., actualización parcial). Instituto de Evaluación Tecnológica en Salud; Instituto Nacional de Cancerología. <http://gpc.minsalud.gov.co>
- Ministerio de Salud y Protección Social. (2014, octubre). *Cáncer de mama, una enfermedad de ascenso en Colombia* (Boletín de prensa). Ministerio de Salud y Protección Social.
- Alejo-Martínez, H., Salazar-Hurtado, E., Poveda-Suárez, C., Puerto-Jiménez, D., Campos, F. R., & Sánchez, O. R. (2014). Impacto del programa de aseguramiento de la calidad de las mamografías en Colombia. *Anales de Radiología México*, 13(4), 369-383.
- Meyblum, E., Gardavaud, F., Dao, T.-H., Fournier, V., Beaussart, P., Pigneur, F., Baranes, L., Rahmouni, A., & Luciani, A. (2015). Breast tomosynthesis: Dosimetry and image quality assessment on phantom. *Diagnostic and Interventional Imaging*, 96(9), 931-939. <https://doi.org/10.1016/j.diii.2014.12.010>
- Guzmán, V. C., Restrepo, H. D. B., & Hurtado, E. S. (2019). Natural scene statistics of mammography accreditation phantom images. In *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)* (pp. 1-5). IEEE. <https://doi.org/10.1109/STSIVA.2019.8730289>

INTELIGENCIA ARTIFICIAL Y APRENDIZAJE FEDERADO PARA LA GESTIÓN INTELIGENTE DE HUERTOS URBANOS

ARTIFICIAL INTELLIGENCE AND FEDERATED LEARNING FOR THE SMART MANAGEMENT OF URBAN GARDENS

Johanna MORENO-CASTELLANOS¹, Laura GRANDE-PÉREZ², Juan M. NÚÑEZ V.³
and Sara RODRÍGUEZ GONZÁLEZ⁴

¹ *Grupo de Investigación BISITE, Universidad de Salamanca. Salamanca, España*
✉ jmorenoc@usal.es
🌐 <https://orcid.org/0009-0008-3309-9891>

² *Grupo de Investigación BISITE, Universidad de Salamanca. Salamanca, España*
✉ lauragrande@usal.es
🌐 <https://orcid.org/0009-0004-2994-8628>

³ *Grupo de Investigación BISITE, Universidad de Salamanca. Salamanca, España*
✉ jmnunez@usal.es
🌐 <https://orcid.org/0000-0002-3165-6200>

⁴ *Grupo de Investigación BISITE, Universidad de Salamanca. Salamanca, España*
✉ srg@usal.es
🌐 <https://orcid.org/0000-0002-3081-5177>

RESUMEN: La seguridad y la soberanía alimentarias en contextos urbanos vulnerables enfrenta limitaciones asociadas a la escasez de recursos, la variabilidad climática y la falta de conectividad. Este trabajo propone una arquitectura inteligente basada en aprendizaje federado (*Federated Learning*) y *edge computing* para la gestión eficiente de huertos urbanos distribuidos. La solución integra sensores IoT que recopilan variables ambientales como temperatura, humedad y luminosidad, permitiendo que cada nodo entrene modelos locales para optimizar la productividad y el uso de recursos hídricos sin compartir datos crudos. Mediante un servidor federador, los modelos locales se agregan utilizando técnicas como FedAvg, generando un modelo global robusto y escalable. La incorporación de inteligencia artificial en el borde reduce la latencia, el consumo de red y mejora la resiliencia ante fallos de conectividad. Los resultados evidencian una mejora en la toma de decisiones, la optimización del riego y la identificación de patrones de consumo, contribuyendo al desarrollo sostenible de la agricultura urbana y a la protección de la privacidad de los datos

Palabras clave: aprendizaje federado; huertos urbanos; edge computing.

ABSTRACT: Food security and sovereignty in vulnerable urban contexts face limitations associated with resource scarcity, climate variability, and lack of connectivity. This work proposes an intelligent architecture based on federated learning and edge computing for the efficient management of distributed urban gardens. The solution integrates IoT sensors that collect environmental variables such as temperature, humidity, and light, allowing each node to train local models to optimize productivity and water resource use without sharing raw data. Through a federating server, local models are aggregated using techniques such as FedAvg, generating a robust and scalable global model. The incorporation of artificial intelligence at the edge reduces latency and network consumption and improves resilience to connectivity failures. The results show an improvement in decision-making, irrigation optimization, and the identification of consumption patterns, contributing to the sustainable development of urban agriculture and the protection of data privacy.

Keywords: federated learning; urban gardens; edge computing.

1. INTRODUCCIÓN

La seguridad y la soberanía alimentarias representan uno de los principales retos contemporáneos [1], especialmente en entornos urbanos con alta vulnerabilidad social y limitaciones de infraestructura. El crecimiento poblacional, el cambio climático y la presión sobre los recursos hídricos han impulsado el desarrollo de enfoques tecnológicos orientados a mejorar la eficiencia y la sostenibilidad de la producción agrícola en pequeña escala. En este contexto, los huertos urbanos se consolidan como una estrategia clave para fortalecer la resiliencia alimentaria local [2], aunque su gestión continúa enfrentando desafíos relacionados con el monitoreo, la optimización de recursos y la toma de decisiones basada en datos.

El estado del arte evidencia una evolución progresiva desde la agricultura inteligente basada en IoT [3] hacia la incorporación de técnicas avanzadas de inteligencia artificial para la predicción y la optimización de variables agrícolas. Estudios recientes destacan el uso de *edge computing* como una solución para reducir la latencia [4], el consumo de ancho de banda y la dependencia de infraestructuras centralizadas, especialmente en escenarios con conectividad limitada. Asimismo, el aprendizaje federado ha emergido como un paradigma prometedor que permite entrenar modelos colaborativos sin necesidad de compartir datos sensibles, garantizando privacidad y escalabilidad [5]. A partir de estas tendencias, este trabajo propone una arquitectura que integra IoT, *edge computing* y aprendizaje federado para la gestión inteligente de huertos urbanos, orientada a optimizar el uso de recursos hídricos, mejorar la productividad y apoyar la seguridad alimentaria en comunidades vulnerables [6].

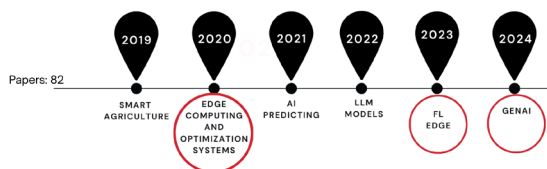


Figura 1. Línea del tiempo del estado del arte. Palabras claves seleccionadas para la investigación

La pregunta de investigación del proyecto es: ¿Cómo se puede contribuir a la seguridad y la soberanía alimentarias de las comunidades vulnerables y no interconectadas mediante la implementación de sistemas de aprendizaje federado agrícola para mejorar la productividad de los huertos urbanos?

2. METODOLOGÍA

La metodología de investigación se desarrolla bajo un enfoque experimental y aplicado, orientado al diseño, la implementación y la validación de una arquitectura inteligente basada en IoT, *edge computing* y aprendizaje federado para la gestión de huertos urbanos. En una primera fase, se realiza el análisis del contexto y el levantamiento de requerimientos, considerando condiciones ambientales, limitaciones de conectividad y necesidades de las comunidades vulnerables donde se implementan los huertos.

Posteriormente, se diseña e implementa la arquitectura IoT-Edge-FL, integrando sensores para la adquisición de variables ambientales como temperatura, humedad y luminosidad, así como nodos de procesamiento en el borde encargados del entrenamiento de modelos locales de aprendizaje automático. Cada nodo entrena su modelo de manera independiente, sin compartir datos crudos, preservando la privacidad de la información.

En la siguiente fase, se implementa el aprendizaje federado, donde los parámetros de los modelos locales son enviados a un servidor federador que consolida la información mediante algoritmos de agregación, como FedAvg, para generar un modelo global. Finalmente, el sistema es validado en huertos urbanos reales o simulados, evaluando métricas de productividad, eficiencia en el uso del agua, latencia y consumo de red, con el fin de analizar el impacto de la propuesta en la toma de decisiones y la sostenibilidad agrícola.



Figura 2. Metodología de investigación

3. RESULTADOS

Los resultados demuestran que la integración de aprendizaje federado y *edge computing* en huertos urbanos mejora la gestión inteligente del cultivo, al reducir la latencia y el consumo de ancho de banda mediante el procesamiento local de datos. Esta característica es especialmente útil en contextos con conectividad limitada. Los modelos locales, entrenados con variables ambientales, optimizan el uso del agua y ajustan el riego de forma más eficiente, mientras que el modelo federado global permite identificar patrones de consumo y productividad para apoyar la toma de decisiones. En conjunto, la solución propuesta contribuye a la sostenibilidad, la privacidad de los datos y la seguridad alimentaria en comunidades vulnerables.

La arquitectura propuesta sigue un enfoque Edge AIoT con aprendizaje federado y se organiza en tres capas: física, Edge y de aplicación. Los sensores IoT recolectan datos ambientales, los dispositivos Edge entrenan modelos locales con baja latencia y la capa de aplicación consolida los parámetros mediante un servidor federador (por ejemplo, FedAvg), generando un modelo global sin compartir datos sensibles. Esta arquitectura escalable y resiliente mejora la productividad agrícola y la optimización de recursos en entornos distribuidos.

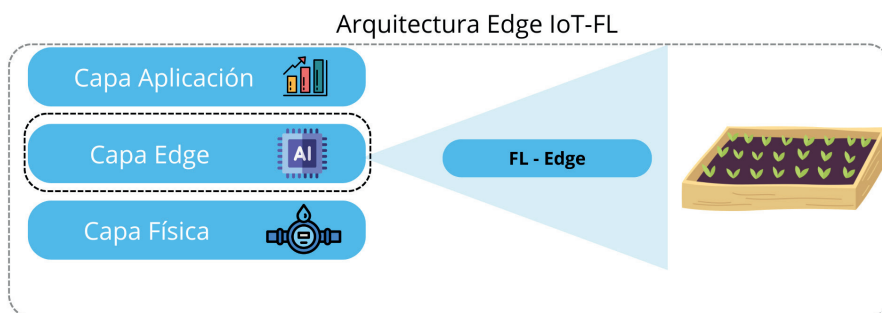


Figura 3. Arquitectura propuesta Edge-FL

4. CONCLUSIONES

La implementación de una arquitectura basada en IoT, *edge computing* y aprendizaje federado demuestra ser una solución efectiva para la gestión inteligente de huertos urbanos, especialmente en comunidades vulnerables y con conectividad limitada. El uso de aprendizaje federado permite entrenar modelos colaborativos sin compartir datos crudos, garantizando la privacidad de la información y promoviendo la soberanía de los datos agrícolas.

Los resultados evidencian que el procesamiento en el borde reduce la latencia y el consumo de red, mejorando la eficiencia operativa y la resiliencia del sistema frente a fallos de conectividad. Asimismo, la optimización del uso de los recursos hídricos y la identificación de patrones de consumo contribuyen a una mayor productividad y sostenibilidad de los cultivos.

Finalmente, esta propuesta tecnológica fortalece la seguridad y la soberanía alimentarias y sienta las bases para el desarrollo de políticas públicas y estrategias de apoyo a la agricultura urbana, integrando la inteligencia artificial como un habilitador clave para el desarrollo social y ambiental.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

6. REFERENCES

1. Grafton, R. Q., Daugbjerg, C., & Qureshi, M. E. (2015). Towards food security by 2050. *Food Security*, 7(2), 179-183. <https://doi.org/10.1007/s12571-015-0445-x>
2. Núñez V., J. M., Corchado, J. M., Giraldo, D. M., Rodríguez-González, S., & De la Prieta, F. (2024). Recommendation system using bio-inspired algorithms for urban orchards. *Internet of Things*, 26, Artículo 101173. <https://doi.org/10.1016/j.iot.2024.101173>
3. Quy, V. K., Hau, N. V., Anh, D. V., Quy, N. M., Ban, N. T., Lanza, S., Randazzo, G., & Muzirafuti, A. (2022). IoT-enabled smart agriculture: Architecture, applications, and challenges. *Applied Sciences*, 12(7), Artículo 3396. <https://doi.org/10.3390/app12073396>
4. El Jarroudi, M., Kouadio, L., Delfosse, P., Bock, C. H., Mahlein, A.-K., Fettweis, X., Mercatoris, B., Adams, F., Lenné, J. M., & Hamdioui, S. (2024). Leveraging edge artificial intelligence for sustainable agriculture. *Nature Sustainability*, 7(7), 846-854. <https://doi.org/10.1038/s41893-024-01352-4>

5. Dembani, R., Karvelas, I., Akbar, N. A., Rizou, S., Tegolo, D., & Fountas, S. (2025). Agricultural data privacy and federated learning: A review of challenges and opportunities. *Computers and Electronics in Agriculture*, 232, Artículo 110048. <https://doi.org/10.1016/j.compag.2025.110048>
6. Núñez V., J. M., Corchado, J. M., & Giraldo, D. M. (2023). Development of an IoT network for urban orchards in high vulnerability areas in Colombia. In *Trends in Sustainable Smart Cities and Territories* (pp. 67-76). Springer. https://doi.org/10.1007/978-3-031-36957-5_7

MATCHING SEMÁNTICO PARA LA REVITALIZACIÓN RURAL: UNA PROPUESTA TECNOLÓGICA ANTE LA DESPOBLACIÓN Y LA EXCLUSIÓN JUVENIL

SEMANTIC MATCHING FOR RURAL REVITALIZATION: A TECHNOLOGICAL PROPOSAL TO ADDRESS DEPOPULATION AND YOUTH EXCLUSION

Natalia L.-GALLEGO¹, Isabella PÉREZ-CAVIEDES² y Danna C. SALAMANCA-VALENZUELA³

¹ *Universidad Autónoma de Occidente. Cali, Colombia*

✉ natalia.lopez_gal@uao.edu.co

🌐 <https://orcid.org/0009-0001-9542-6631>

² *Universidad Autónoma de Occidente. Cali, Colombia*

✉ isabella.perez_cav@uao.edu.co

🌐 <https://orcid.org/0009-0009-4213-3534>

³ *Universidad Autónoma de Occidente. Cali, Colombia*

✉ danna.salamanca@uao.edu.co

🌐 <https://orcid.org/0009-0006-6053-2932>

RESUMEN: El despoblamiento rural y el desequilibrio generacional constituyen retos estructurales persistentes, mientras una proporción significativa de jóvenes permanece excluida de los sistemas de educación, empleo o formación (NEET). Este trabajo presenta Ruralink, una propuesta tecnológica para la reconexión entre juventud y territorio mediante un sistema de *matching* semántico explicable, concebido desde los principios de la inteligencia artificial fiable, la gobernanza de datos y el control humano. La contribución no reside en un nuevo algoritmo, sino en demostrar cómo técnicas consolidadas de procesamiento del lenguaje natural pueden integrarse de forma responsable en contextos de política pública, combinando representaciones semánticas, reglas territoriales, explicabilidad y trazabilidad institucional. La plataforma integra modelos *Sentence-Transformers*, un esquema híbrido de puntuación y un asistente conversacional orientado a reducir barreras informativas y burocráticas, manteniendo protección de datos y supervisión humana. Se describen la arquitectura, el *pipeline* operacional y un marco de evaluación orientado a impacto territorial, aportando un modelo replicable de intermediación sociotécnica alineado con DemIA para abordar simultáneamente la exclusión juvenil y el reto demográfico.

Palabras clave: *matching* semántico; revitalización rural; *embeddings* semánticos; despoblación rural; exclusión juvenil neet; inteligencia artificial; PNL.

ABSTRACT: Rural depopulation and generational imbalance are persistent structural challenges, while a significant proportion of young people remain excluded from education, employment, or training (NEET). This paper presents Ruralink, a technological proposal for reconnecting young people and their territory through an explainable semantic matching system, designed based on the principles of reliable artificial intelligence, data governance, and human control. The contribution lies not in a new algorithm, but in demonstrating how established natural language processing techniques can be responsibly integrated into public policy contexts, combining semantic representations, territorial rules, explainability, and institutional traceability. The platform integrates Sentence-Transformers models, a hybrid scoring scheme, and a conversational assistant aimed at reducing informational and bureaucratic barriers, while maintaining data protection and human oversight. The architecture, operational pipeline, and an evaluation framework focused on territorial impact are described, providing a replicable model of socio-technical intermediation aligned with DemIA to simultaneously address youth exclusion and the demographic challenge.

Keywords: semantic matching; rural revitalization; semantic embeddings; rural depopulation; NEET youth exclusion; artificial intelligence; NLP.

1. INTRODUCCIÓN

La despoblación rural y la concentración demográfica en áreas urbanas constituyen uno de los principales desafíos socioeconómicos contemporáneos, mientras una proporción relevante de jóvenes permanece excluida de los sistemas formales de educación y empleo (NEET). Esta coexistencia de territorios con déficit demográfico y juventud excluida evidencia una desconexión estructural entre oportunidades territoriales, aspiraciones individuales y mecanismos de intermediación.

En este contexto, la intermediación inteligente puede desempeñar un papel clave al traducir perfiles y restricciones en oportunidades viables; sin embargo, cuando incorpora inteligencia artificial en ámbitos sensibles como el empleo o la movilidad territorial, resulta imprescindible garantizar principios de explicabilidad, transparencia, gobernanza de datos y control humano, en línea con la Estrategia Nacional de Inteligencia Artificial (ENIA) y los objetivos de DemIA. Este trabajo propone *Ruralink* como un marco de integración de técnicas de procesamiento de lenguaje natural (PLN) explicables para la intermediación territorial, orientado a la revitalización territorial y la inclusión juvenil.

2. ANTECEDENTES Y ESTADO DEL ARTE

España y América Latina enfrentan una tendencia sostenida de despoblación rural, con casi la mitad de las provincias españolas en riesgo de perder población en los próximos 20

años [2, 6]. Esta crisis se entrelaza con la exclusión de los jóvenes NEET, quienes representan el 16,5 % de la población de entre 15 y 29 años debido a la precarización urbana y la falta de recursos para el relevo generacional. Aunque iniciativas como Proyecto Arraigo y la Red Nacional de Pueblos Acogedores han validado metodologías exitosas de relocalización y activación de servicios mediante acompañamiento humano, ambas presentan retos críticos de escalabilidad y automatización en sus procesos de seguimiento [4, 5].

La incorporación de inteligencia artificial permite optimizar este emparejamiento mediante el uso de *embeddings* contextuales y transformadores que facilitan el análisis de similitud coseno y *clustering* para vincular perfiles con oportunidades [7, 8]. Estos motores de recomendación superan el filtrado tradicional al permitir una personalización que anticipa necesidades territoriales [9, 10]. Asimismo, la integración de *chatbots* LLM transforma la interacción en la plataforma al automatizar consultas de gestión territorial, siempre bajo un marco de gobernanza ética y protección de datos personales. La evidencia internacional confirma que combinar la personalización algorítmica con la participación comunitaria mejora sustancialmente el acceso a servicios públicos y amplía el alcance de las políticas de revitalización rural [11, 12].

3. PROPUESTA METODOLÓGICA: RURALINK

El objetivo de Ruralink es facilitar emparejamientos útiles, comprensibles y verificables entre jóvenes y oportunidades territoriales, priorizando la explicabilidad del proceso de recomendación, la protección de datos personales y la evaluación del impacto social. El sistema está diseñado como una herramienta de apoyo a la toma de decisiones, no como un mecanismo automatizado de asignación.

3.1. ARQUITECTURA

La arquitectura sigue un enfoque modular y desacoplado, orientado a escalabilidad y reproducibilidad. El sistema se compone de:

- 1) un *frontend web* con formularios guiados y visualización de resultados;
- 2) un *frontend web* con formularios guiados y visualización de resultados;
- 3) una API de *matching* implementada en Python con Fast API;
- 4) una capa de datos gobernados sobre Supabase, que integra autenticación, almacenamiento relacional y políticas de Row-Level Security; y un asistente conversacional que interactúa con los usuarios y accede de forma controlada a los servicios inteligentes.

Este diseño permite una adopción progresiva por parte de administraciones locales y organizaciones sociales, sin requerir infraestructuras complejas.

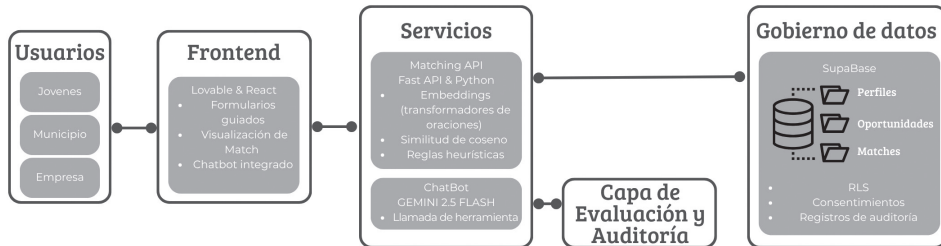


Figura 1. Arquitectura lógica del sistema de *matching* semántico para la revitalización territorial

3.2. PIPELINE DE MATCHING SEMÁNTICO EXPLICABLE

El proceso de emparejamiento se estructura en las siguientes etapas:

- 1) Normalización del perfil ciudadano y oportunidades, combinando campos estructurados y descripciones textuales.
- 2) Generación de *embeddings* semánticos mediante modelos Sentence-Transformers.
- 3) Cálculo de similitud entre perfiles y oportunidades usando similitud coseno.
- 4) *Scoring* híbrido, que ajusta el ranking semántico incorporando requisitos críticos, compatibilidad territorial y criterios económicos.
- 5) Generación de explicaciones, produciendo razones legibles que justifican cada recomendación.
- 6) Persistencia y retroalimentación, almacenando resultados y *feedback* para análisis longitudinal y mejora iterativa.

3.3. GOBERNANZA DE DATOS E IA FIABLE

Ruralink implementa privacidad por diseño mediante consentimientos granulares, auditoría de operaciones y mecanismos de revocación. La trazabilidad del proceso de *matching* y la generación explícita de explicaciones facilitan la supervisión humana y la evaluación de sesgos, alineando la solución con los principios de inteligencia artificial fiable promovidos por DemIA.

4. RESULTADOS Y DISCUSIÓN

Dado el carácter piloto de la propuesta, este trabajo adopta una evaluación prospectiva, orientada a definir criterios de validación técnica e impacto social en futuras

implementaciones. Se proponen métricas técnicas como latencia del proceso, estabilidad del ranking y tasa de aceptación de recomendaciones, junto con indicadores de impacto territorial a medio plazo, incluyendo inserción laboral, permanencia en el territorio y satisfacción de los usuarios. Desde una perspectiva crítica, la calidad de los emparejamientos depende tanto de la completitud de los perfiles como de la sostenibilidad real de las oportunidades ofertadas.

5. CONCLUSIONES Y TRABAJO FUTURO

La plataforma *Ruralink* propone un marco tecnológico y organizativo para enfrentar la exclusión juvenil y el reto demográfico mediante *matching* semántico explicable. Su aporte central es mostrar cómo técnicas de NLP pueden aplicarse de forma responsable en política pública, priorizando explicabilidad, gobernanza de datos y control humano sobre la optimización algorítmica.

El trabajo demuestra que la IA, concebida como apoyo y no como mecanismo opaco de decisión, fortalece la legitimidad y la efectividad de las políticas territoriales. Como proyección, se plantea el pilotaje en territorios concretos, la recolección sistemática de retroalimentación institucional y la evaluación longitudinal del impacto social y demográfico, con miras a consolidar un modelo replicable de intermediación sociotécnica alineado con DemIA.

6. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

7. REFERENCIAS

1. Collantes, F., & Pinilla, V. (2019). La despoblación rural en España: causas, impactos y respuestas. *Revista Española de Estudios Rurales*, 22, 33-50.
2. Red Española de Desarrollo Rural. (s.f.). *Informe sobre despoblación y desarrollo rural en España*. <https://www.redrural.es>
3. Proyecto Neets In Action. (2020). Informe técnico sobre fragmentación de programas juveniles.
4. Proyecto Arraigo. (s.f.). *Repoblación: Impacto y metodología*. <https://www.proyectoarraigo.es/repoblacion/>

5. Pueblos Acogedores. (s.f.). Red Nacional de Pueblos Acogedores. <https://pueblosacogedores.com>
6. Schlögel, K. (2022). Repensando el territorio rural europeo: innovación y declive demográfico. *Otero Rural*, 21(2), 39-66.
7. Sanabria-Medina, G. (2024). Aplicaciones de Sentence Transformers en embeddings semánticos [Manuscrito no publicado].
8. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need*. *Advances in Neural Information Processing Systems*, 30, 5998-6008.
9. Sanabria-Medina, G., & Regil-Vargas, J. (2023). Inteligencia artificial para la recomendación de recursos en educación a distancia. *Revista de Educación Artificial*, 10(2), 933-948.
10. Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2011). *Recommender systems handbook*. Springer. <https://doi.org/10.1007/978-0-387-85820-3>
11. Chen, M., Wang, L., & Zhang, H. (2024). Impulsando la revitalización rural: liberando el potencial de la economía digital. *Journal of Rural Revitalization*, 9(2), 35-61.
12. Sánchez-Céspedes, J. M. (2022). Aplicación de la inteligencia artificial en la formulación de políticas públicas rurales. *Revista SciELO Colombia*, 40, 111-139.

PLATAFORMA ACCESIBLE PARA ALERTAS TEMPRANAS HIDROMETEOROLÓGICAS BASADA EN PRONÓSTICOS CLIMÁTICOS Y UMBRALES LOCALES DE PRECIPITACIÓN

ACCESSIBLE PLATFORM FOR HYDROMETEOROLOGICAL EARLY WARNINGS BASED ON CLIMATE FORECASTS AND LOCAL PRECIPITATION THRESHOLDS

Camila A. CARDONA-ALZATE¹ y David Alejandro CAJIAO-LAZT²

¹ *Universidad Autónoma de Occidente. Cali, Colombia*

✉ camila_and.cardona@uao.edu.co

🌐 <https://orcid.org/0009-0003-3201-9187>

² *Universidad Autónoma de Occidente. Cali, Colombia*

✉ david.cajiao@uao.edu.co

🌐 <https://orcid.org/0009-0000-5992-7079>

RESUMEN: La información climática sigue siendo inaccesible para las comunidades vulnerables en Latinoamérica, a pesar de la existencia de sistemas abiertos de pronóstico. Este artículo presenta una plataforma que aprovecha los principios de la predicción basada en impactos al traducir los pronósticos de precipitación en alertas tempranas para desastres naturales inducidos por lluvias en zonas rurales de Colombia. El objetivo es aportar una plataforma colaborativa para usuarios con baja alfabetización. El proyecto busca alertar aprovechando los umbrales de precipitación en combinación con pronósticos climáticos para anticipar condiciones de lluvia riesgosas. Utilizando una metodología *data thinking*, se desarrolló un prototipo alpha. Aprovechando herramientas climáticas y la evaluación de riesgos, la plataforma muestra cómo las soluciones accesibles e inclusivas pueden cerrar la brecha entre los servicios climáticos globales y la preparación local para desastres, promoviendo los ODS 11 y 13.

Palabras clave: resiliencia climática; IBF; Impact-Based Forecasting; comunidades vulnerables, Open-Meteo.

ABSTRACT: Climate information remains inaccessible to vulnerable communities in Latin America, despite the existence of open forecasting systems. This article presents a platform that leverages the principles of impact-based prediction by translating precipitation forecasts

into early warnings for rain-induced natural disasters in rural areas of Colombia. The goal is to provide a collaborative platform for users with low literacy levels. The project seeks to provide alerts by using precipitation thresholds in combination with climate forecasts to anticipate risky rainfall conditions. Using a data thinking methodology, an alpha prototype was developed. Leveraging climate tools and risk assessment, the platform shows how accessible and inclusive solutions can bridge the gap between global climate services and local disaster preparedness, promoting SDGs 11 and 13.

Keywords: climate resilience; IBF; impact-based forecasting; vulnerable communities, open-meteo.

1. INTRODUCCIÓN

Colombia enfrenta un incremento dramático en la frecuencia y la intensidad de fenómenos hidrometeorológicos extremos: entre 2001 y 2021 se registraron 23.722 eventos (58 % inundaciones, 34,3 % movimientos en masa) [1], y actualmente 308 municipios están bajo alerta por deslizamientos, 68 en nivel rojo [2]. Las comunidades más vulnerables carecen de acceso a información climática en tiempo real, enfrentan barreras de alfabetización digital y conectividad limitada. La OMM ha promovido IBFWS (Servicios de Pronóstico y Alerta Basados en Impactos) para transformar pronósticos en información accionable [3], pero su operacionalización en contextos de bajos recursos permanece como un desafío no resuelto. A pesar de sistemas abiertos disponibles (Open-Meteo, CHIRPS, NOAA), existe desconexión crítica entre servicios climáticos globales y comunidades locales vulnerables. Este trabajo contribuye con: (i) arquitectura IBFWS accesible de bajo costo basada en API abiertas; (ii) prototipo de alta fidelidad integrando umbrales de precipitación con pronósticos en tiempo real e interfaces inclusivas; (iii) metodología de diseño *data-driven* con retroalimentación comunitaria; (iv) integración de ciencia climática rigurosa con datos históricos oficiales y conocimiento local; (v) contribución conceptual a ODS 11 y 13 demostrando cómo IBFWS con accesibilidad fortalecen anticipación comunitaria ante desastres climáticos.

2. ESTADO DEL ARTE

2.1. SERVICIOS DE PRONÓSTICO Y ALERTA BASADOS EN IMPACTOS (IBFWS)

Los IBFWS (Impact-Based Forecast and Warning Services) son promovidos por la OMM como un enfoque que integra pronóstico del peligro con información de vulnerabilidad/exposición y evaluación de riesgo con el fin de convertir pronósticos complejos en información accionable para decisiones preventivas [3]. En aplicaciones humanitarias, la Cruz Roja/Media Luna Roja ha mostrado que la anticipación basada en pronósticos puede habilitar acciones tempranas (p. ej., distribución de ayuda antes del evento usando disparadores) y reducir pérdidas, como se ha documentado en Bangladesh [4]. A escala

regional, el SAHF resalta la urgencia de pasar de servicios meteorológicos tradicionales a IBFWS en múltiples horizontes de anticipación (de horas a décadas), apoyándose en capacitación, codesarrollo con sectores (agricultura, agua, salud, gestión del riesgo) y uso de pronósticos por conjuntos [5]. En una investigación reciente, Kooshki Forooshani *et al.* proponen un modelo de ML (XGBoost) para predecir impactos de ciclones tropicales (porcentaje de viviendas totalmente dañadas) a escala municipal, mostrando que el uso de datos globales abiertos puede mantener un desempeño comparable, aunque se requiere validación en otros países para confirmar transferibilidad [6].

2.2. UMBRALES DE PRECIPITACIÓN

En umbrales de precipitación para alerta temprana, la literatura enfatiza que la lluvia actúa como disparador crítico de desastres en regiones montañosas tropicales y que los umbrales deben distinguir entre lluvia desencadenante (1-6 h de alta intensidad) y lluvia preparatoria (24-72 h acumulada que aumenta la humedad del suelo y reduce la resistencia) [7]. La revisión sistemática de González *et al.* también identifica enfoques físicos y estadísticos, proponiendo su integración como práctica deseable para robustecer la definición de umbrales [7]. Dou *et al.* plantean un *framework* de umbrales con índice de lluvia dinámico e integración con susceptibilidad espacial, permitiendo que los umbrales se adapten a condiciones hidrometeorológicas variables y articulen explícitamente la dimensión espacial (dónde) y temporal (cuándo) del peligro [8]. En Colombia, Marín y Marín-Sánchez presentan LandScient_EWS, *software open-source* que compara lluvia medida en tiempo real con umbrales intensidad-duración tipo *power-law*, calcula el Intensity Ratio (IR) y genera niveles de alerta operacionales en diferentes escalas. Su validación incluye un caso de estudio en Medellín (deslizamiento de Villatina, 2008) [9].

3. METODOLOGÍA

Se implementó un enfoque de investigación conceptual centrado en datos integrando análisis de requisitos, diseño inclusivo y prototipado de mediana fidelidad. Se caracterizó el problema en contexto colombiano (alta exposición a eventos extremos, baja conectividad y alfabetización digital) y se definieron objetivos estratégicos con OKR cuantificables. Se levantaron requisitos funcionales (interfaz intuitiva, multicanalidad app/SMS, confiabilidad, multilingüismo) y especificaciones técnicas (API climáticas, nube escalable, bajo consumo de datos). Mediante ideación se generaron dos propuestas (A: integración multicanal con retroalimentación comunitaria; B: acción rápida con interfaz mínima), ambas prototipadas en mediana fidelidad. Se aplicó análisis jerárquico de procesos (AHP) con criterios ponderados (precisión 0,51, acciones comunitarias 0,24, rapidez 0,17, usabilidad 0,08), seleccionando Propuesta A (0,63 vs. 0,37). La arquitectura final comprende cuatro capas: datos (Open-Meteo), procesamiento (navegador del usuario, sin servidor), aplicación (interfaz web con

integración futura SMS) y accesibilidad (lenguaje simplificado, iconografía, audio, bajo consumo de datos).

4. RESULTADOS

El pipeline de datos opera en el navegador del usuario con arquitectura de cuatro capas: adquisición (Open-Meteo API, resolución $0,1^\circ \times 0,1^\circ$, actualización cada 6 horas; geolocalización del dispositivo), procesamiento (clasificación de códigos WMO a español; sistema de alertas contra umbrales predefinidos: Peligro >10 mm precipitación/ >50 km/h viento/WMO ≥ 95 , Alerta >5 mm/ >30 km/h/WMO ≥ 80 , Advertencia >0 mm/ >20 km/h/WMO ≥ 60 , calibrados según lluvia desencadenante [4]), aplicación (visualización geoespacial con Google Maps, módulo de preparación con guías narradas en audio y *checklist* de emergencia, funcionalidad de reporte ciudadano geolocalizados) y accesibilidad (lenguaje simplificado, iconografía, audio, bajo consumo de datos).

La implementación utiliza React con Vite, componentes accesibles (shadcn/ui, Radix UI, Lucide React), Tailwind CSS para diseño responsivo y TypeScript para tipado estático. El sistema consume cuatro servicios externos: Open-Meteo (pronósticos), Nominatim OSM (geocodificación inversa), Geolocation API (coordenadas, *fallback*: Cali $3,4516^\circ\text{N}$ $76,5320^\circ\text{W}$) y FormSubmit.co (envío de reportes). La Figura 1 ilustra la conexión de API externas, herramientas de procesamiento local de umbrales y servicios de infraestructura (Docker, Google Cloud, Nginx). El despliegue utiliza Docker en dos etapas: compilación con Node.js/Vite y producción con Nginx Alpine sirviendo *assets* estáticos, permitiendo despliegue escalable sin *backend* dedicado. Las métricas de rendimiento registraron tiempo de construcción de 80 segundos y tiempo de respuesta inicial (TTFB) de 618 milisegundos, asegurando carga rápida incluso con conectividad limitada.

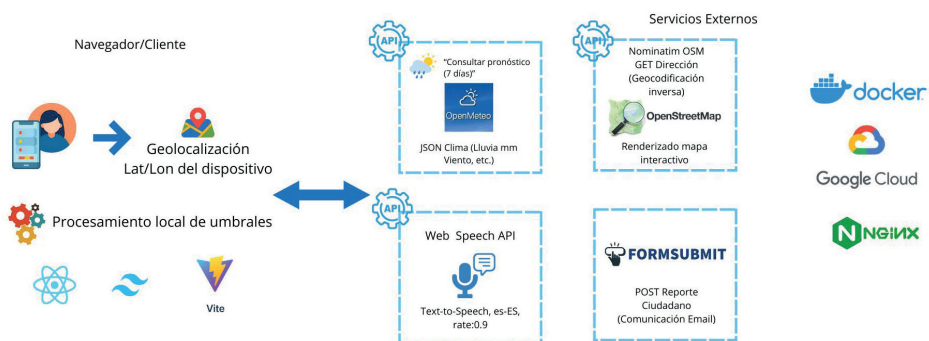


Figura 1. Pipeline de datos y experiencia de usuario

5. CONCLUSIONES

Este trabajo demuestra la viabilidad conceptual de un servicio IBFWS accesible para contextos vulnerables latinoamericanos, integrando pronósticos climáticos abiertos, umbrales de precipitación calibrados localmente y diseño inclusivo. Las limitaciones actuales incluyen ausencia de validación con usuarios finales, umbrales calibrados solo para análisis general requiriendo validación regional específica y proyecciones de impacto teóricas sin comportamiento real documentado. Las próximas fases contemplan despliegue piloto en municipios rurales colombianos, calibración regional de umbrales y evaluación rigurosa de impacto.

6. DOCUMENTACIÓN

Este proyecto se podrá encontrar con su documentación en este link de Github: <https://github.com/DCajiao/kani-weather-app-alert>

7. REFERENCIAS

1. Fondo de Población de las Naciones Unidas - Colombia. (2023). Riesgo de desastres y cambio climático. En *ASP Colombia 2023* (Cap. 3.9). UNFPA. https://colombia.unfpa.org/sites/default/files/pub-pdf/3.9_riesgo_de_desastres.pdf
2. Instituto de Hidrología, Meteorología y Estudios Ambientales. (2025). *Pronóstico de la amenaza por deslizamientos de tierra* (Boletín de Alerta por Pronóstico de la Amenaza N.º 210). IDEAM. https://www.ideam.gov.co/sites/default/files/prensa/boletines/2025-07-29/210_badt_julio_29_2025.pdf
3. World Meteorological Organization. (2021). *Guidelines on multi-hazard impact-based forecast and warning services. Part II: Putting multi-hazard IBFWS into practice* (WMO-No. 1150). WMO. <https://library.wmo.int/records/item/57739-wmo-guidelines-on-multi-hazard-impact-based-forecast-and-warning-services>
4. Red Cross Red Crescent Climate Centre, & UK Met Office. (2021). *The future of forecasts: Impact-based forecasting for early action* (3rd ed.). Anticipation Hub. https://www.anticipation-hub.org/Documents/Manuals_and_Guidelines/RCCC_Impact_based_forecasting_Guide_2021-3.pdf
5. Regional Integrated Multi-hazard Early Warning System. (2023). *Impact based forecasting* (South Asia Hydromet Forum, SAHF Working Paper 3). RIMES. https://www.sahf.info/wp-content/uploads/2023/03/SAHF_Working-Paper_3_Impact-based-Forecasting.pdf
6. Kooshki Forooshani, M., van den Homberg, M. J. C., Kalimeri, K., Kaltenbrunner, A., Mejova, Y., Milano, L., Ndirangu, P., Paolotti, D., Teklesadik, A., & Turner, M. L. (2024). Towards a global impact-based forecasting model for tropical cyclones. *Natural Hazards and Earth System Sciences*, 24(1), 309-329. <https://doi.org/10.5194/nhess-24-309-2024>
7. Gonçalves Gonzalez, F. C., do C. Reis Cavalcanti, M., Nahas Ribeiro, W., Barreto de Mendonça, M., & Naked Haddad, A. (2023). A systematic review on rainfall thresholds for landslides occurrence. *Heliyon*, 10(1), Artículo e23247. <https://doi.org/10.1016/j.heliyon.2023.e23247>

8. Dou, H., Chang, J., Zhao, B., Wang, G., Wang, Y., & Liu, S. (2023). Rainfall early warning threshold and its spatial distribution under dynamic rainfall index. *International Journal of Applied Earth Observation and Geoinformation*, 116, Artículo 103195. <https://doi.org/10.1016/j.jag.2023.103195>
9. Marín, R. J., & Marín-Sánchez, J. C. (2024). LandScient_EWS: Real-time monitoring of rainfall thresholds for landslide early warning - A case study in the Colombian Andes. *Journal of Engineering Geology*, 34(2), 173-191. <https://doi.org/10.9720/kseg.2024.2.173>

TEMPORAL FUSION TRANSFORMERS PARA LA
PREDICCIÓN DEL TÉRMINO DE FACTURACIÓN PVPC

TEMPORAL FUSION TRANSFORMERS (TFT) FOR THE
PREDICTION OF THE PVPC BILLING TERM

Víctor GARCÍA-CHICO¹, Manuel RAMOS-JIMÉNEZ², Lucía RIERA-PRESMANES³, Roberto SALAMANCA-BARRIOS⁴, y Ana Belén GIL GONZÁLEZ⁵

¹ Universidad de Salamanca. Salamanca, Castilla y León, España

✉ victorgarciachico@usal.es

🌐 <https://orcid.org/0009-0001-4834-2697>

² Universidad de Salamanca. Salamanca, Castilla y León, España

✉ manuelramosji@usal.es

🌐 <https://orcid.org/0009-0006-2852-1766>

³ Universidad de Salamanca. Salamanca, Castilla y León, España

✉ luripres@usal.es

🌐 <https://orcid.org/0009-0006-7254-6311>

⁴ Universidad de Salamanca. Salamanca, Castilla y León, España

✉ robertosala@usal.es

🌐 <https://orcid.org/0009-0007-2318-6054>

⁵ Universidad de Salamanca. Salamanca, Castilla y León, España

✉ anabg@usal.es

🌐 <https://orcid.org/0000-0001-7235-6151>

RESUMEN: La creciente volatilidad del mercado eléctrico, impulsada por la penetración de fuentes renovables, exige modelos de predicción de precios (EPF) precisos para garantizar una gestión energética eficiente en el sector residencial. Este trabajo aborda la predicción a corto plazo (horizonte 24 horas) de la tarifa regulada española (PVPC) mediante la arquitectura *Temporal Fusion Transformer* (TFT), diseñada para capturar dependencias temporales complejas manteniendo la interpretabilidad. Utilizando datos reales de 2024, se evalúa el desempeño del modelo propuesto frente a redes LSTM y métodos de persistencia. Los resultados experimentales demuestran la superioridad del TFT, logrando un MAE de 30.63 €/MWh y un sMAPE de 0.1838, lo que representa una reducción del error del 28.7 % respecto a la arquitectura LSTM convencional.

Palabras clave: predicción de precios de electricidad (EPF); PVPC; Temporal Fusion Transformer (TFT); *deep learning*; series temporales; gestión energética.

ABSTRACT: The increasing volatility of the electricity market, driven by the penetration of renewable energy sources, requires accurate Electricity Price Forecasting (EPF) models to ensure efficient energy management in the residential sector. This work addresses short-term forecasting (24-hour horizon) of the Spanish regulated tariff (PVPC) using the Temporal Fusion Transformer (TFT) architecture, designed to capture complex temporal dependencies while maintaining interpretability. Using real data from 2024, the performance of the proposed model is evaluated against LSTM networks and persistence-based methods. The experimental results demonstrate the superiority of the TFT, achieving an MAE of 30.63 €/MWh and an SMAPE of 0.1838, representing a 28.7 % reduction in error compared to the conventional LSTM architecture.

Keywords: electricity price forecasting (EPF); PVPC; temporal fusion transformer (TFT); *deep learning*; time series; energy management.

1. INTRODUCCIÓN

La descarbonización del sector eléctrico ha convertido la predicción de precios (EPF) en una competencia crítica para los sistemas de gestión energética del hogar (HEMS), que deben anticipar la alta volatilidad introducida por las renovables para optimizar el consumo [2]. En España, este desafío es central para la tarifa regulada (PVPC), indexada directamente al mercado mayorista.

Si bien el estado del arte en EPF ha evolucionado desde modelos econométricos hacia arquitecturas de *Deep Learning* (como LSTM), persisten limitaciones significativas: los métodos clásicos no capturan la complejidad no lineal del mix energético [4], mientras que las redes neuronales operan comúnmente como «cajas negras», careciendo de la explicabilidad necesaria para la toma de decisiones [1].

Para abordar esta brecha, este trabajo propone la aplicación del *Temporal Fusion Transformer* (TFT) [3] al mercado español. A diferencia de los modelos recurrentes estándar, el TFT emplea mecanismos de atención para aprender dependencias temporales a largo plazo manteniendo una interpretabilidad inherente. El objetivo de este estudio es validar mediante un primer modelo TFT si esta arquitectura presenta cierto potencial para superar en precisión a los modelos de referencia (LSTM).

2. DATOS Y METODOLOGÍA

2.1. CONJUNTO DE DATOS

Para este estudio se han utilizado datos reales del mercado eléctrico español obtenidos a través de la API de ESIOS (Red Eléctrica de España). La variable objetivo es el término

de facturación del PVPC con impuestos incluidos (*PVPC_billing_term_with_taxes*), ya que representa el coste final real para el consumidor doméstico. El *dataset* abarca registros horarios históricos desde junio de 2021 hasta diciembre de 2024. Se ha realizado una partición cronológica estricta para evitar el filtrado de información futura (*data leakage*):

- *Entrenamiento*: Del 1 de junio de 2021 hasta el 30 de junio de 2024.
- *Validación*: Del 1 de julio al 30 de septiembre de 2024 (3 meses).
- *Test*: Del 1 de octubre al 31 de diciembre de 2024 (3 meses).

Las variables de entrada se dividieron en dos categorías según su disponibilidad en tiempo real:

1. Conocidas en el futuro (*Known Inputs*): Variables de calendario (hora, día de la semana, mes, festivos nacionales) y tipos impositivos (IVA, IEE).
2. Observadas en el pasado (*Unknown Inputs*): Histórico de precios, tarifas de acceso (periodos valle/llano/pico), demanda real y generación por tecnologías (eólica, solar, nuclear, ciclo combinado). Estas variables solo están disponibles para el codificador (pasado) y no para el decodificador (futuro).

2.2. MODELOS Y CONFIGURACIÓN EXPERIMENTAL

Se han evaluado tres enfoques para la predicción con un horizonte de 24 horas (*Day-Ahead*):

1. *Baseline* de persistencia: Modelo ingenuo que asume que el perfil de precios de mañana será idéntico al de hoy: ($P_{t+24} = P_t$). Sirve como cota inferior de rendimiento.
2. *Long Short-Term Memory (LSTM)*: Red recurrente clásica configurada con 2 capas ocultas de 64 unidades y *dropout* del 0.1. Se entrenó minimizando el error absoluto medio (MAE). A diferencia del TFT, esta arquitectura trata todas las entradas de manera secuencial sin distinguir explícitamente entre variables estáticas o futuras conocidas.
3. *Temporal Fusion Transformer (TFT)*: La arquitectura propuesta se configuró con un tamaño de estado oculto de 64, 1 cabezal de atención y *QuantileLoss* para estimar la incertidumbre. El modelo utiliza una ventana de contexto (*encoder length*) de 168 horas (7 días) para predecir las siguientes 24 horas. El entrenamiento se realizó utilizando el optimizador Ranger y *Early Stopping* para prevenir el sobreajuste.

3. RESULTADOS

La Tabla 1 resume el desempeño de los modelos en el conjunto de test (último trimestre de 2024). El modelo TFT superó significativamente tanto al *baseline* como a la red LSTM en ambas métricas evaluadas.

Tabla 1. Comparativa de métricas de error en el conjunto de Test

Modelo	MAE (€MWh)	sMAPE (%)
Baseline (Persistencia)	38.10	32.19 %
LSTM (Recurrente)	43.00	25.65 %
TFT (Propuesto)	30.63	18.38 %

Mientras que la LSTM logró reducir el error porcentual (sMAPE) respecto a la persistencia, su error absoluto (MAE) fue mayor, indicando dificultades para manejar los picos de precios o cambios de tendencia abruptos. Por el contrario, el TFT logró una reducción del 28.7 % en MAE respecto a la LSTM y del 19.6 % respecto al *baseline*.

La Figura 1 muestra una predicción de 24 horas del modelo TFT frente a la realidad y al modelo de persistencia. Se observa cómo el modelo es capaz de anticipar correctamente la estructura diaria de precios, ajustándose a las rampas de subida y bajada mejor que los modelos tradicionales, gracias a la incorporación efectiva de las variables de calendario futuras.

4. CONCLUSIONES Y TRABAJO FUTURO

Este estudio preliminar valida la eficacia del *Temporal Fusion Transformer* para la predicción de precios en el mercado eléctrico español regulado (PVPC). Los resultados demuestran que el TFT no solo supera a las arquitecturas recurrentes clásicas (LSTM) en precisión (sMAPE 0.18 vs 0.25), sino que ofrece una arquitectura más robusta para integrar variables heterogéneas (series temporales pasadas y variables de calendario futuras).

La superioridad del TFT sugiere que los mecanismos de atención son más efectivos que la recurrencia simple para capturar las dependencias temporales complejas del mercado eléctrico actual, altamente influenciado por la volatilidad de las renovables.

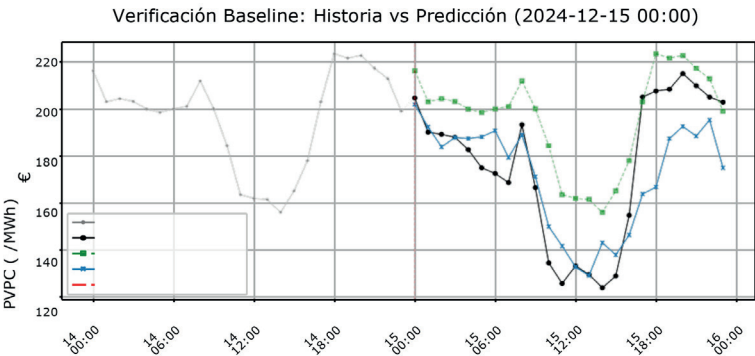


Figura 1. Predicción de precios a 24h (TFT vs Realidad y *Baseline*) 15-12-2024, 00:00

Este trabajo representa una primera aproximación hacia un sistema de predicción avanzado. Gracias a que los TFT permiten incorporar *inputs* futuros conocidos (ej. demanda programada) y precios en contextos determinados, se abre una vía de investigación para desarrollar modelos que mejoren drásticamente la precisión de las redes neuronales tradicionales.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

6. REFERENCIAS

1. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
2. Lago, J., Marcjasz, G., De Neyer, B., & Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Renewable and Sustainable Energy Reviews*, 133, Artículo 110317. <https://doi.org/10.1016/j.rser.2020.110317>
3. Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
4. Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4), 1030-1081. <https://doi.org/10.1016/j.ijforecast.2014.08.008>

PREDICCIÓN DEL RIESGO DE INCENDIOS FORESTALES MEDIANTE TELEDETECCIÓN E IA (2020-2025)

FOREST FIRE RISK PREDICTION USING REMOTE SENSING AND ARTIFICIAL INTELLIGENCE

Rodrigo ÁLVAREZ¹, Javier SANTOS², Thide E. LLORENTE³, Stephany NIEVES⁴
y Samuel REY⁵

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ ralvarezmartin@usal.es

🌐 <https://orcid.org/0009-0004-1157-4373>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ javier.santos@usal.es

🌐 <https://orcid.org/0009-0008-4310-4746>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ llorentel@usal.es

🌐 <https://orcid.org/0000-0001-6543-0026>

⁴ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ snievesu@usal.es

🌐 <https://orcid.org/0009-0003-6937-2072>

⁵ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ srey@usal.es

🌐 <https://orcid.org/0009-0002-4675-5065>

RESUMEN: Se sintetizan los avances de 2020-2025 en predicción del riesgo de incendios forestales mediante teledetección y modelos de inteligencia artificial. Se resume el uso de índices espectrales (NDVI, EVI, NDMI), variables térmicas, topográficas y antrópicas, y se comparan líneas base (regresión, Random Forest, SVM) con arquitecturas profundas (CNN, ConvLSTM, Transformers). Se destacan resultados recientes con AUC entre 0.85 y 0.96 y F1 entre 0.67 y 0.85, así como la incorporación de autosupervisión y adaptación eficiente de parámetros para transferir modelos entre regiones. Se proponen buenas prácticas de evaluación (F1/AUC, RMSE/MAE y Kappa) y recomendaciones para despliegue operativo con datos multifuente y restricciones de costo computacional.

Palabras clave: wildfire risk; teledetección; CNN; ConvLSTM; Transformers; LoRA.

ABSTRACT: Advances from 2020-2025 in forest fire risk prediction using remote sensing and artificial intelligence are synthesized. The use of spectral indices (NDVI, EVI, NDMI), thermal, topographic, and anthropogenic variables is summarized, and baseline approaches (regression, Random Forest, SVM) are compared with deep learning architectures (CNN, ConvLSTM, Transformers). Recent results reporting AUC values between 0.85 and 0.96 and F1 scores between 0.67 and 0.85 are highlighted, along with the incorporation of self-supervised learning and parameter-efficient fine-tuning techniques to transfer models across regions. Best practices for evaluation (F1/AUC, RMSE/MAE, and Kappa) and recommendations for operational deployment using multi-source data under computational cost constraints are also proposed.

Keywords: wildfire risk; remote sensing; CNN; ConvLSTM; Transformers; LoRA.

1. INTRODUCCIÓN

Los incendios forestales amenazan ecosistemas, economías y salud pública. Entre 2019 y 2023 los grandes incendios en Australia y Canadá quemaron millones de hectáreas y generaron pérdidas superiores a 100 000 millones de dólares, mientras que en España la superficie quemada en 2022 (~247,667 ha) quintuplicó la del año previo. El incremento simultáneo de temperaturas, sequías y presión antrópica refuerza la necesidad de mapas de riesgo más precisos para orientar prevención y preposicionamiento de recursos.

La literatura reciente combina datos satelitales ópticos y térmicos, reanálisis meteorológicos y variables socioeconómicas para estimar probabilidades de ignición y propagación. Revisiones recientes destacan la integración de datos multifuente y modelos profundos para mejorar la predicción en diferentes regiones [4]. Este artículo condensa los hallazgos clave 2020-2025 sobre fuentes de datos, modelos de predicción y métricas de evaluación, con foco en la aplicabilidad operativa y la transferibilidad espacio-temporal.

2. DATOS SATELITALES Y VARIABLES DE ENTRADA

2.1. ÍNDICES ESPECTRALES Y HUMEDAD

El NDVI continúa como indicador estándar de vigor de la vegetación, aunque sufre saturación en coberturas densas y sensibilidad al fondo del suelo [1]. El EVI añade la banda azul para corregir aerosoles y mejorar la respuesta en bosques densos [2]. El SAVI incorpora un factor de suelo y el NDMI combina NIR-SWIR para estimar el contenido hídrico de la vegetación, con valores negativos asociados a suelo desnudo y positivos altos a vegetación sin estrés [3].

Estos índices se complementan con NDWI y NBR para humedad y áreas quemadas y con temperatura superficial (LST) para detectar estrés térmico previo a la ignición. Los productos de área quemada como MCD64A1 (MODIS, colección 6) sirven para calibrar

y validar mapas de riesgo [10]. El preprocesamiento incluye enmascarado de nubes, armonización radiométrica y remuestreo a celdas comunes; las series diarias o semanales suelen suavizarse con ventanas móviles para reducir ruido. En regiones con alta nubosidad se recurre a interpolaciones y a la fusión con radar de apertura sintética para mantener continuidad en estimaciones de humedad del combustible [11].

2.2. TOPOGRAFÍA, USOS DEL SUELO Y METEOROLOGÍA

Pendiente, orientación y altitud condicionan la acumulación de combustible y la dirección de propagación; la proximidad a carreteras y áreas urbanas explica la mayor parte de igniciones de origen humano. Variables meteorológicas (temperatura del aire, humedad relativa, precipitación, viento) se integran mediante reanálisis como ERA5 y el índice Fire Weather Index, mejorando la predicción diaria y subestacional del riesgo [9].

La combinación de capas multifuente demanda normalización espacial y temporal para evitar sesgos por resolución. La inclusión de indicadores socioeconómicos (densidad de población, usos del suelo) ha demostrado incrementar la sensibilidad del modelo en interfaces urbano-forestales sin degradar la especificidad en áreas remotas.

Para reducir el desequilibrio entre celdas con y sin incendio, se emplean muestreos estratificados y ponderaciones temporales que priorizan días cercanos a eventos. Los conjuntos se dividen respetando la independencia espacial (por cuencas o cuadrículas) para evitar fugas de información entre entrenamiento y prueba.

3. MODELOS DE PREDICCIÓN

3.1. MODELOS TRADICIONALES

Las regresiones logísticas y los árboles de decisión fueron las primeras líneas base; permiten interpretar el efecto de variables como pendiente o NDVI, pero capturan de forma limitada umbrales no lineales. Random Forest y SVM mejoran la precisión en conjuntos heterogéneos y siguen siendo útiles como referencias rápidas y explicables [4].

No obstante, la autocorrelación espacial y la dinámica temporal reducen su capacidad predictiva frente a métodos que modelan patrones espaciales y secuencias. Por ello se emplean como línea de base o para generar mapas explicativos que acompañen a modelos más complejos.

Variantes geográficamente ponderadas y modelos de mezcla permiten cierta adaptación local, pero su rendimiento decrece en escenarios multiespectrales de alta dimensión. El valor principal de estas líneas base reside en su interpretabilidad y en su uso para auditorías y detección de sesgos.

3.2. CNN Y ARQUITECTURAS HÍBRIDAS

Las CNN explotan la estructura espacial de los mapas de riesgo. Un estudio con 11 507 incendios históricos en Guangdong (China) reportó en test precisión del 89.7 %, *recall* del 90.7 %, F1 = 84.9 % y AUC = 0.962, superando a árboles de decisión, RF y regresión logística [5]. Las variables más influyentes fueron humedad relativa, temperatura, NDVI y precipitación.

Variantes tipo ResNet y U-Net preservan detalles finos y se combinan con *autoencoders* o módulos LSTM para segmentación semántica del riesgo. Estos enfoques mejoran la detección de focos dispersos y reducen falsas alarmas en paisajes heterogéneos, manteniendo inferencias en segundos por escena.

La interpretabilidad se apoya en mapas de calor (Grad-CAM) que resaltan zonas críticas y en estimadores de incertidumbre (por ejemplo, dropout Monte Carlo) para priorizar revisiones humanas. Ensamblajes ligeros de CNN incrementan la robustez sin penalizar severamente el tiempo de inferencia.

3.3. MODELOS ESPACIO-TEMPORALES Y *TRANSFORMERS*

Las ConvLSTM mantienen la coherencia espacial en secuencias multitemporales y han mostrado mejoras frente a LSTM y XGBoost en escenarios mediterráneos [4]. A escala global, redes híbridas ConvLSTM alcanzan AUC cercanas a 0.89 y F1 alrededor de 0.72 al equilibrar memoria temporal y precisión espacial. *Transformers* geoespaciales capturan dependencias de largo alcance y correlaciones climáticas; variantes deformables y con atención en ventanas han superado a CNN en bosques de pino al integrar variables meteorológicas y combustibles [4]. En detección en tiempo casi real, enfoques autosupervisados como SIT-FUSE distinguen frentes de fuego y columnas de humo con SSIM de 0.71-0.86 [6]. La adaptación eficiente de parámetros mediante LoRA permite afinar modelos de fundación geoespacial usando solo 5-10 % de los parámetros, facilitando su traslado entre regiones [7].

La mayoría de estos modelos se entrenan con validación cruzada espacio-temporal y ventanas deslizantes que respetan la causalidad. La adaptación de dominio incorpora capas de ajuste o pérdidas adversarias para reducir el corrimiento entre regiones; los mejores resultados combinan preentrenamiento autosupervisado, *fine-tuning* eficiente y datos locales de alta resolución.

4. EVALUACIÓN Y MÉTRICAS

Las métricas principales son precisión, *recall*, F1 y área bajo la curva ROC (AUC). Los mejores resultados recientes con CNN se sitúan en AUC cercanas a 0.96 y F1 alrededor de 0.85 [5], mientras que modelos espacio-temporales se mantienen en AUC cercanas a 0.89. En mapas categóricos, el coeficiente Kappa y su versión ponderada cuantifican el acuerdo espacial más allá del azar [8].

Se recomienda evaluar un vector de métricas: F1/AUC para discriminación, RMSE/MAE para errores absolutos en probabilidades o superficie quemada, y Kappa para acuerdo

espacial. Para despliegue operativo deben añadirse tiempo de inferencia y consumo energético, priorizando modelos eficientes cuando el refresco diario sea crítico.

La partición de datos debe evitar fugas: separar por bloques espaciales o periodos completos y reservar años recientes como conjunto de prueba. En contextos desbalanceados, las curvas *precision-recall* y el costo de falsas alarmas por región ofrecen una lectura más informativa que la ROC agregada.

5. RETOS Y OPORTUNIDADES FUTURAS

Interpretabilidad y confianza siguen siendo críticas: técnicas SHAP y LIME permiten derivar umbrales de cobertura o humedad donde el riesgo se dispare, como mostró el proyecto Wildfire Genome al asociar coberturas de coníferas y altitud con incrementos bruscos de riesgo [8]. La transferibilidad espacial requiere adaptaciones ligeras (LoRA, adaptadores) para ajustar modelos preentrenados a nuevas regiones con pocos datos [7].

La fusión de datos multiresolución (ópticos, térmicos, radar, socioeconómicos) demanda normalización robusta y puede beneficiarse de autosupervisión para explotar grandes volúmenes sin etiquetar [6, 11]. Reducir el costo computacional mediante compresión y actualización parcial de parámetros es clave para que los modelos profundos sean sostenibles en centros operativos y dispositivos de borde.

La gobernanza de datos (versionado, trazabilidad de capas imputadas y control de acceso) y la estandarización de formatos facilitan la replicabilidad entre agencias. La integración de LiDAR para estimar carga de combustible tridimensional complementa las señales ópticas y radar y mejora la estimación de la humedad en la subcopia [12].

6. CONCLUSIONES

Los índices de vegetación y humedad continúan siendo la base de los mapas de riesgo, pero su integración con variables topográficas, meteorológicas y antrópicas incrementa la sensibilidad y la especificidad. Las CNN y ConvLSTM superan de forma consistente a regresiones, RF y SVM, con AUC entre 0.85 y 0.96 y F1 superiores a 0.7 [4, 5], mientras que *Transformers* y enfoques autosupervisados amplían el horizonte de predicción y reducen la dependencia de etiquetas [6].

Las métricas combinadas (F1/AUC, RMSE/MAE, Kappa) y el reporte de tiempos de inferencia deben acompañar cualquier despliegue operativo. La adaptación eficiente de parámetros habilita la transferencia entre regiones sin reentrenar desde cero, acercando modelos de alta precisión a escenarios reales con limitaciones de datos y recursos [7, 9].

7. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia

Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

8. REFERENCIAS

1. NASA Earthdata. (s.f.). *Normalized Difference Vegetation Index (NDVI)*. Earth Science Data Systems, NASA. <https://www.earthdata.nasa.gov/topics/land-surface/normalized-difference-vegetation-index-ndvi>
2. U.S. Geological Survey. (s.f.). *Landsat Enhanced Vegetation Index (EVI)*. U.S. Department of the Interior. <https://www.usgs.gov/landsat-missions/landsat-enhanced-vegetation-index>
3. Geovisualization.net. (2022, 14 de julio). *Measuring moisture – Normalized Difference Moisture Index (NDMI) Sentinel-2, 2022*. <https://geovisualization.net/2022/07/14/measuring-moisture-normalized-difference-moisture-index-ndmi-sentinel-2-2022/>
4. Xu, Z., Wang, X., & Zhang, Y. (2024). *Wildfire risk prediction: A review* (arXiv:2405.01607). arXiv. <https://arxiv.org/abs/2405.01607>
5. Jiang, W., Wang, L., Chen, H., & Liu, J. (2024). Wildfire risk assessment using deep learning in Guangdong Province, China. *International Journal of Applied Earth Observation and Geoinformation*, 128, Artículo 103750. <https://doi.org/10.1016/j.jag.2024.103750>
6. LaHaye, N., El-Kadi, H., & Garay, M. (2025). Harnessing self-supervised deep learning and geostationary remote sensing for advancing wildfire and associated air quality monitoring (SIT-FUSE). In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2025)*. IEEE. <https://doi.org/10.1109/IGARSS55030.2025.11243034>
7. Thoreau, R., Lefèvre, S., & Courty, N. (2025). Parameter-efficient adaptation of geospatial foundation models through embedding deflection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2025)*. IEEE. <https://doi.org/10.1109/ICCV51701.2025.00895>
8. Liu, C., & Mostafavi, A. (2025). *WildfireGenome: Interpretable machine learning reveals local drivers of wildfire risk and their cross-county variation* (arXiv:2511.11589). arXiv. <https://arxiv.org/abs/2511.11589>
9. Vitolo, C., Di Giuseppe, F., Krzeminski, B., & San-Miguel-Ayanz, J. (2020). ERA5-based global meteorological wildfire danger maps. *Scientific Data*, 7, Artículo 214. <https://doi.org/10.1038/s41597-020-0554-z>
10. Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, 217, 72-85. <https://doi.org/10.1016/j.rse.2018.08.005>
11. Chrysafis, I., Mallinis, G., Gitas, I., & Tsakiri-Strati, M. (2023). Vegetation fuel mapping using Sentinel-1, Sentinel-2, and DEM derivatives—East Macedonia and Thrace, Greece. *Remote Sensing*, 15(4), Artículo 1015. <https://doi.org/10.3390/rs15041015>
12. Bright, B. C., Hudak, A. T., Meddens, A. J. H., Hawbaker, T. J., Briggs, J., & Kennedy, R. E. (2022). Multitemporal lidar captures heterogeneity in fuel loads and consumption on the Kaibab Plateau. *Fire Ecology*, 18, Artículo 16. <https://doi.org/10.1186/s42408-022-00142-7>

SEIR-XAI: INFERENCIA EXPLICABLE DE LA TRANSMISIÓN EPIDÉMICA PARA ANÁLISIS DE ESCENARIOS Y APOYO A DECISIONES

SEIR-XAI: EXPLAINABLE INFERENCE OF EPIDEMIC TRANSMISSION FOR SCENARIO ANALYSIS AND DECISION SUPPORT

Marco HERNÁNDEZ PÉREZ¹, Pablo-Enrique GUILLEM FERNÁNDEZ² y Javier PRIETO³

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ marcohper@usal.es

🌐 <https://orcid.org/0000-0002-6381-5149>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ pabloe.guifer@usal.es

🌐 <https://orcid.org/0009-0001-5328-7661>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ javierp@usal.es

🌐 <https://orcid.org/0000-0001-8175-2201>

RESUMEN: Los modelos epidemiológicos clásicos SEIR ofrecen interpretabilidad, pero presentan una flexibilidad limitada para capturar factores contextuales dinámicos, mientras que los enfoques de aprendizaje automático suelen carecer de transparencia. Presentamos SEIR-XAI, un marco híbrido que integra un modelo SEIR en tiempo discreto con aprendizaje automático explicable para inferir la tasa de transmisión efectiva β como una variable dependiente del contexto. Mediante modelos interpretables y explicaciones basadas en SHAP, SEIR-XAI permite un análisis tanto global como local de cómo las intervenciones políticas y los factores ambientales influyen en la dinámica de transmisión. La evaluación con datos de COVID-19 muestra una mejora en el ajuste del modelo en comparación con un modelo SEIR base, manteniendo al mismo tiempo la interpretabilidad. Los resultados ponen de relieve que una modelización epidémica eficaz no debe centrarse únicamente en la precisión predictiva, sino también en proporcionar información transparente y accionable para el apoyo a la toma de decisiones.

Palabras clave: modelización epidemiológica; SEIR; inteligencia artificial explicable; apoyo a la toma de decisiones.

ABSTRACT: Classical SEIR epidemiological models offer interpretability but limited flexibility to capture dynamic contextual factors, while machine learning approaches often lack transparency. We present SEIR-XAI, a hybrid framework that integrates a discrete-time SEIR model with explainable machine learning to infer the effective transmission rate β as a context-dependent variable. Using interpretable models and SHAP-based explanations, SEIR-XAI enables both global and local analysis of how policy interventions and environmental factors influence transmission dynamics. Evaluation on COVID-19 data shows improved model fit compared to a baseline SEIR model while preserving interpretability. The results highlight that effective epidemic modeling should not only aim at predictive accuracy, but also at providing transparent and actionable insights for decision support.

Keywords: epidemiological modeling; SEIR; explainable AI; decision support.

1. INTRODUCCIÓN Y MOTIVACIÓN

Los modelos compartimentales de tipo SIR y SEIR han sido durante décadas la base de la modelización epidemiológica, gracias a su formulación mecanicista y a la interpretación directa de sus parámetros [1]. Sin embargo, su aplicación a escenarios reales complejos se ve limitada por supuestos de homogeneidad y por la dificultad de incorporar factores contextuales dinámicos, como políticas públicas o cambios de comportamiento [2]. Durante la pandemia de COVID-19, estas limitaciones impulsaron el uso de enfoques basados en aprendizaje automático, que mejoran el ajuste a datos observados, pero introducen problemas de interpretabilidad y confianza, especialmente relevantes en contextos de decisión de alto impacto [3]. En este trabajo se parte de la hipótesis de que es posible reforzar los modelos epidemiológicos clásicos sin renunciar a su interpretabilidad, mediante la inferencia dinámica y explicable de parámetros clave.

Presentamos *SEIR-XAI*, un sistema que mantiene la estructura compartimental SEIR e introduce una estimación explicable y dependiente del contexto de la tasa de transmisión β . El objetivo no es sustituir el modelo epidemiológico, sino mejorar su capacidad para informar decisiones mediante explicaciones transparentes y cuantificables.

2. MARCO METODOLÓGICO: MODELO SEIR-XAI

El núcleo del sistema es un modelo SEIR formulado en tiempo discreto, adecuado para trabajar con datos epidemiológicos diarios o semanales. La dinámica del sistema se describe mediante los compartimentos Susceptible (S), Expuesto (E), Infectado (I) y Recuperado (R), parametrizados por la tasa de transmisión β , el periodo de incubación y el periodo infeccioso. A diferencia de enfoques clásicos que asumen β constante, en SEIR-XAI este parámetro se considera dinámico y dependiente del contexto, permitiendo capturar cambios en la transmisión asociados a intervenciones no farmacológicas y condiciones externas sin modificar la estructura del modelo [4].

2.1. INFERENCIA DINÁMICA Y EXPLICABLE DE β

La estimación de β se plantea como un problema de regresión supervisada utilizando ventanas temporales deslizantes. Para cada ventana, β se infiere a partir de variables contextuales observadas, que incluyen indicadores de políticas públicas (índice de dureza gubernamental, cierre de escuelas y confinamiento), factores ambientales y contexto temporal. Se emplea un ensamble de modelos explicables, combinando modelos aditivos generalizados y Gradient Boosting. La interpretabilidad se garantiza mediante SHAP [5], que permite descomponer cada estimación de β en contribuciones individuales de las variables de entrada.

3. RESULTADOS Y ANÁLISIS EXPLICABLE

El rendimiento de SEIR-XAI se comparó con un modelo SEIR basal adaptativo utilizando métricas estándar de ajuste (Tabla 1).

Tabla 1. Comparación de rendimiento entre modelos

Modelo	R ²	MAE	RMSE
SEIR basal adaptativo	0.42	0.118	0.156
SEIR-XAI	0.74	0.061	0.083

SEIR-XAI muestra una mejora clara en todas las métricas, evidenciando que la inferencia dinámica de β permite reproducir de forma más precisa la dinámica epidémica observada.

El análisis de importancia global basado en valores SHAP indica que las variables asociadas a intervenciones no farmacológicas explican la mayor parte de la variabilidad de β . El índice de dureza gubernamental con un retardo de 14 días presenta la mayor contribución media absoluta ($|\text{SHAP}| \approx 0.087$), seguido por el cierre de escuelas ($|\text{SHAP}| \approx 0.021$) y las órdenes de confinamiento ($|\text{SHAP}| \approx 0.011$). La temperatura muestra una contribución menor, pero consistente ($|\text{SHAP}| \approx 0.009$), mientras que el resto de variables presentan impacto marginal.

Las explicaciones locales permiten descomponer estimaciones individuales de β en contribuciones de las variables de entrada. En determinadas ventanas temporales, incrementos en la dureza gubernamental se asocian a contribuciones positivas a β , reflejando la dinámica temporal entre la implementación de medidas y la transmisión observada. La evolución temporal de β estimada por SEIR-XAI reproduce de forma coherente las distintas fases de la pandemia y presenta una mayor concordancia con los valores observados que el modelo SEIR basal adaptativo, consistente con la mejora observada en las métricas de ajuste.

4. DISCUSIÓN Y CONCLUSIONES

SEIR-XAI constituye una alternativa intermedia entre los modelos epidemiológicos clásicos y los enfoques de aprendizaje automático opacos. Frente al SEIR clásico, permite incorporar contexto dinámico sin alterar la estructura compartimental; frente a modelos de caja negra, proporciona explicaciones trazables tanto a nivel global como local.

El sistema se encuentra actualmente validado en condiciones controladas (TRL4). Asimismo, las explicaciones obtenidas reflejan asociaciones derivadas de datos observacionales y no relaciones causales estrictas, lo que requiere una interpretación cuidadosa.

En conclusión, el principal valor de SEIR-XAI no reside únicamente en predecir mejor, sino en *informar mejor*. El enfoque demuestra que es posible reforzar modelos epidemiológicos clásicos mediante aprendizaje automático explicable para generar conocimiento accionable, transparente y útil para el apoyo a la toma de decisiones.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto RESILIENCE: REcurSos médicos mediante InteLIgENCIA artificial explicable. Referencia: CNS2022-135101. Entidad financiadora: Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033) y European Union NextGenerationEU/PRTR.

6. REFERENCIAS

- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 700-721. <https://doi.org/10.1098/rspa.1927.0118>
- Brauer, F., Castillo-Chavez, C., & Feng, Z. (2019). *Mathematical models in epidemiology* (Vol. 32). Springer. <https://doi.org/10.1007/978-1-4939-9828-9>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Majumdar, S. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4), 529-538. <https://doi.org/10.1038/s41562-021-01079-8>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.

SISTEMA DE APRENDIZAJE ADAPTATIVO CON IA PARA UNIVERSITARIOS

AI-BASED ADAPTIVE LEARNING SYSTEM FOR UNIVERSITY STUDENTS

Santiago MURGUEITIO MUÑOZ¹ y Juan Manuel NÚÑEZ VELASCO²

¹ *Universidad Autónoma de Occidente. Cali, Colombia*

✉ santiago_mur@uao.edu.co

🌐 <https://orcid.org/0009-0000-6691-2537>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jmnunez@usal.es

🌐 <https://orcid.org/0000-0002-3165-6200>

RESUMEN: La educación superior enfrenta el desafío de atender una creciente diversidad de ritmos, necesidades y trayectorias de aprendizaje en un contexto de acelerada transformación tecnológica. La irrupción de herramientas basadas en inteligencia artificial generativa (IA Gen) abre oportunidades para fortalecer los procesos formativos, pero también plantea interrogantes sobre su uso responsable y su impacto pedagógico. Este artículo presenta el diseño y el desarrollo de un prototipo de alta fidelidad de un ecosistema de aprendizaje adaptativo apoyado por IA generativa, orientado a complementar —y no reemplazar— la labor docente. El sistema integra recursos académicos curados por profesores, representados mediante *embeddings* almacenados en una base de datos vectorial, con información contextual del estudiante (desempeño, habilidades e intereses) para habilitar agentes conversacionales personalizados por asignatura. El enfoque pedagógico del agente es mayéutico, privilegiando la guía y la reflexión por sobre la entrega directa de respuestas. El proyecto se encuentra en etapa de desarrollo y avanzará hacia fases de prueba y validación. La propuesta contribuye a la discusión sobre el uso ético y responsable de la IA en educación superior y se alinea con el Objetivo de Desarrollo Sostenible 4 (Educación de calidad).

Palabras clave: IA generativa; aprendizaje adaptativo; educación superior; sistemas tutoriales inteligentes; agentes.

ABSTRACT: Higher education faces the challenge of addressing an increasingly diverse range of learning paces, needs, and trajectories in a context of rapid technological

transformation. The emergence of generative artificial intelligence (GenAI) tools creates opportunities to strengthen educational processes, while also raising questions about responsible use and pedagogical impact. This article presents the design and development of a high-fidelity prototype of an adaptive learning ecosystem supported by generative AI, aimed at complementing —rather than replacing— teaching activities. The system integrates faculty-curated academic resources, represented through embeddings stored in a vector database, with contextual student information (performance, skills, and interests) to enable subject-specific personalized conversational agents. The pedagogical approach of the agent is maieutic, prioritizing guidance and reflection over the direct provision of answers. The project is currently in the development stage and will advance toward testing and validation phases. The proposal contributes to the discussion on the ethical and responsible use of AI in higher education and aligns with Sustainable Development Goal 4 (Quality Education).

Keywords: generative AI; adaptive learning; higher education; intelligent tutoring systems; agents.

1. INTRODUCCIÓN

La educación superior enfrenta una transformación impulsada por la IA generativa, que amplía las posibilidades de apoyo pedagógico, pero también intensifica los debates éticos y pedagógicos frente a la diversidad de ritmos de aprendizaje y las limitaciones de los modelos educativos tradicionales [1, 2, 4, 7, 9]. La rigidez curricular y las restricciones de tiempo del profesorado dificultan el acompañamiento personalizado, favoreciendo un uso no guiado de la tecnología que puede debilitar el aprendizaje profundo y generar dependencias improductivas [2, 3, 4, 6]. En este contexto, el proyecto propone integrar la IA de forma responsable como un complemento a la labor docente, mediante un ecosistema de aprendizaje adaptativo orientado a promover experiencias más inclusivas, fortalecer la autonomía del estudiante y favorecer un aprendizaje reflexivo y contextualizado [2, 4, 7, 8, 9].

2. ESTADO DEL ARTE

Los sistemas de tutoría inteligente (ITS) han evolucionado desde reglas rígidas hacia arquitecturas híbridas que buscan replicar la tutoría humana mediante el modelado del aprendizaje [1, 9]. Como se observa en la Figura 1, esta área mantuvo un crecimiento constante pero moderado durante décadas, enfocándose en ofrecer instrucción personalizada comparable a la de un tutor humano [2]. No obstante, históricamente la investigación se ha concentrado en pruebas de viabilidad y estudios de caso, evidenciando una brecha persistente en evaluaciones longitudinales robustas y una implementación completa que demuestre impacto pedagógico a largo plazo [2, 10].

Un punto de inflexión crítico ocurre en 2017 con la arquitectura de atención (*Transformers*) [11], hito que marca el inicio de una aceleración en la producción científica

visible en la tendencia ascendente del gráfico. Esta innovación permitió el desarrollo de modelos de lenguaje de gran escala (LLMs), cuya adopción masiva tras el lanzamiento de ChatGPT a finales de 2022 [4] disparó el volumen de publicaciones mostrado en la Figura 1. Este fenómeno ha transformado los ITS tradicionales en agentes conversacionales generativos, permitiendo una personalización del lenguaje antes inalcanzable, aunque la evidencia empírica sobre su efectividad educativa aún se encuentra en una fase incipiente y fragmentada [5, 7].

A pesar de este auge tecnológico, la literatura advierte sobre riesgos como el cognitive offloading, la reducción del esfuerzo crítico y la dependencia del estudiante [3, 6]. Por ello, las tendencias actuales —reflejadas en el pico máximo de publicaciones recientes del gráfico— se orientan hacia diseños que integran supervisión humana (human-in-the-loop) y técnicas de recuperación contextualizada (RAG) para mitigar alucinaciones [8, 13]. En este escenario, el presente proyecto se sitúa en la convergencia de los ITS históricos y la IA generativa, utilizando bases vectoriales [13] y agentes mayéuticos para transformar la actual explosión de datos en una herramienta pedagógica controlada y académicamente relevante.

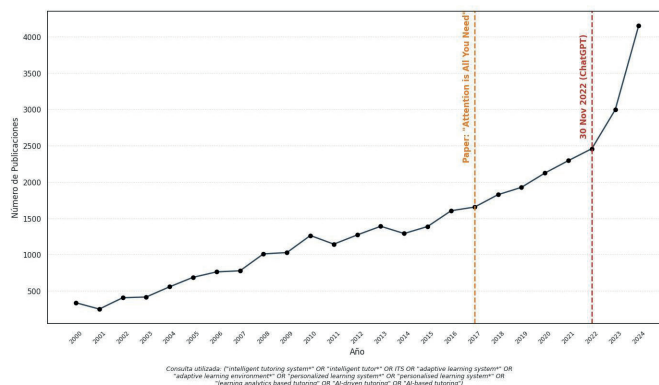


Figura 1. Tendencia de Publicaciones Académicas (Scopus)

3. METODOLOGÍA

El proyecto se desarrolló siguiendo la metodología Data Thinking, la cual enfatiza la comprensión profunda del problema antes de la selección de soluciones tecnológicas, así como el uso de los datos como eje central para el diseño y la toma de decisiones [10]. En una primera fase se realizó el análisis del problema educativo y la definición de objetivos. Posteriormente, se diseñó un prototipo funcional que permitiera evaluar la viabilidad conceptual del ecosistema propuesto.

Se definieron dos roles principales: profesor y estudiante. El profesor puede acceder a sus asignaturas, gestionar contenidos y cargar recursos académicos (libros, textos, materiales

complementarios), replicando procesos habituales de los campus virtuales universitarios. Estos recursos son procesados y transformados en representaciones vectoriales (*embeddings*), que se almacenan en una base de datos vectorial.

4. RESULTADOS PRELIMINARES

El resultado principal de esta etapa es un prototipo de alta fidelidad funcional que demuestra la viabilidad técnica y conceptual del enfoque propuesto. El sistema permite la creación de experiencias de acompañamiento personalizadas y la recomendación de recursos alineados con los intereses y necesidades de cada estudiante. Dado que el proyecto se encuentra en fase de desarrollo, aún no se cuenta con resultados empíricos derivados de pruebas con usuarios finales.

5. CONCLUSIONES

El proyecto demuestra la viabilidad técnica y pedagógica de un sistema de aprendizaje adaptativo basado en inteligencia artificial generativa para la educación superior. La propuesta integra la curaduría de contenidos por parte del docente con mecanismos de personalización centrados en el estudiante, mediante agentes conversacionales por asignatura apoyados en *embeddings* de materiales académicos y datos de desempeño. El diseño mayéutico del agente busca promover la reflexión, la autonomía y el pensamiento crítico, mitigando los riesgos de dependencia cognitiva asociados al uso acrítico de la IA generativa. Si bien los resultados corresponden a una etapa de diseño y prototipado, el sistema se perfila como un complemento al acompañamiento docente y no como un sustituto, alineándose con enfoques de uso ético y responsable de la tecnología. Futuras fases deberán centrarse en la validación empírica con estudiantes y profesores para evaluar su impacto real en el aprendizaje y la experiencia formativa.

6. REFERENCIAS

1. Contrino, M. F., Reyes-Millán, M., Vázquez-Villegas, P., & Membrillo-Hernández, J. (2024). Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education for a more personalized approach. *International Journal of Educational Technology in Higher Education*, 21, Artículo 29. <https://doi.org/10.1186/s40561-024-00292-y>
2. Du Plooy, M., de la Harpe, R., & van Lingen, J. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(19), Artículo e39630. <https://doi.org/10.1016/j.heliyon.2024.e39630>
3. Gerlich, M. (2024). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), Artículo 6. <https://doi.org/10.3390/soc15010006>

4. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Mewes, T., Claudia, P., Radutskiy, P., Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Artículo 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
5. Laun, M., & Wolff, F. (2025). Chatbots in education: Hype or help? A meta-analysis. *Learning and Individual Differences*, 102, Artículo 102646. <https://doi.org/10.1016/j.lindif.2025.102646>
6. Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM. <https://doi.org/10.1145/3706598.3713778>
7. McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2024). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, 88, 1-21. <https://doi.org/10.1007/s10734-024-01288-w>
8. Tarun, B., Du, H., Kannan, D., & Gehringer, E. F. (2025). Human-in-the-loop systems for adaptive learning using generative AI. In *Proceedings of the Frontiers in Education Conference (FIE 2025)*. IEEE. <https://doi.org/10.1109/FIE63693.2025.11328658>
9. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Artículo 39. <https://doi.org/10.1186/s41239-019-0171-0>
10. Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data-analytic thinking and data mining*. O'Reilly Media.
11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need* (arXiv:1706.03762). arXiv. <https://doi.org/10.48550/arXiv.1706.03762>
12. ScienceDirect. (2024). *Intelligent tutoring systems: Topic overview and evolution*. Elsevier.
13. Pinecone. (2024). *Vector database*. <https://www.pinecone.io/>

SISTEMAS DE TUTORÍA INTELIGENTE Y NARRATIVAS DIGITALES PARA PRESERVAR LA DIVERSIDAD BIOCULTURAL DE ECUADOR

INTELLIGENT TUTORING SYSTEMS AND DIGITAL NARRATIVES TO PRESERVE ECUADOR'S BIOCULTURAL DIVERSITY

Juan Pablo LÓPEZ GOYEZ¹, Alfonso GONZÁLEZ BRIONES² y Jesús ARANGUREN³

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jplopezg@usal.es

 <https://orcid.org/0000-0003-2873-2185>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ alfonsogb@usal.es

 <https://orcid.org/0000-0002-3444-4393>

³ *Universidad Politécnica Estatal del Carchi. Antisana, S/N, 040101, Tulcán, Ecuador*

✉ jesus.aranguren@upec.edu.ec

 <https://orcid.org/0000-0002-6989-1097>

RESUMEN: Este trabajo destaca la importancia de valorar la diversidad biocultural, que integra la biodiversidad de los territorios y los saberes, prácticas y tradiciones de los pueblos indígenas, configurando un patrimonio que actualmente se encuentra amenazado por procesos de globalización, homogeneización cultural y despoblación rural. En Ecuador, las comunidades enfrentan una progresiva pérdida de sus conocimientos ancestrales, lenguas y prácticas, como consecuencia de factores económicos y sociales, que debilitan su transmisión intergeneracional. Los sistemas de tutoría inteligente emergen como herramientas para la preservación y la revitalización del patrimonio biocultural. Estos sistemas permiten la creación de narrativas digitales construidas a partir de relatos orales, testimonios y documentos históricos, facilitando procesos de aprendizaje contextualizados. Se propone una arquitectura MAS para el diseño de un sistema de tutoría biocultural, que integra materiales y contenidos generativos orientados a la valoración de los saberes ancestrales. Este enfoque busca revitalizar la identidad cultural y lingüística, y contribuir a la conservación de la memoria colectiva de las comunidades, donde el agente actúa como mediador pedagógico y sociocultural.

Palabras clave: biocultural; MAS; narrativas digitales; pueblos indígenas; ITS.

ABSTRACT: This work highlights the importance of valuing biocultural diversity, which integrates territorial biodiversity with the knowledge, practices, and traditions of Indigenous peoples, shaping a heritage that is currently threatened by processes of globalization, cultural homogenization, and rural depopulation. In Ecuador, communities face a progressive loss of ancestral knowledge, languages, and practices due to economic and social factors that weaken intergenerational transmission. Intelligent Tutoring Systems emerge as tools for the preservation and revitalization of biocultural heritage. These systems enable the creation of digital narratives built from oral histories, testimonies, and historical documents, facilitating contextualized learning processes. A multi-agent system (MAS) architecture is proposed for the design of a biocultural tutoring system that integrates generative materials and content aimed at fostering appreciation of ancestral knowledge. This approach seeks to revitalize cultural and linguistic identity and contribute to the preservation of the collective memory of communities, where the agent acts as both a pedagogical and sociocultural mediator.

Keywords: biocultural; MAS (multi-agent systems); digital narratives; indigenous peoples; intelligent tutoring systems (ITS).

1. INTRODUCCIÓN

La diversidad biocultural constituye uno de los patrimonios más significativos para los territorios latinoamericanos, al integrar la riqueza ecológica con los saberes, las prácticas y las cosmovisiones de los pueblos originarios. En Ecuador, los pueblos y las nacionalidades indígenas mantienen un legado cultural transmitido a través de relatos orales y prácticas comunitarias que articulan identidad, territorio y espiritualidad. La acelerada globalización, la homogenización cultural y la desigualdad estructural amenazan la continuidad de estos saberes, debilitando la memoria colectiva [1].

La educación emerge para revitalizar los conocimientos ancestrales y promover la cohesión territorial. Las investigaciones recientes subrayan la importancia de incorporar narrativas digitales y estrategias de mediación cultural que permitan conectar las epistemologías locales con experiencias contextualizadas. La integración pedagógica de las narrativas y la mediación tecnológica favorece procesos significativos de aprendizaje que reconocen la diversidad [2, 3].

Aún persisten brechas que limitan la posibilidad de adaptar los recursos educativos a las realidades culturales de las comunidades. La literatura en sistemas de tutoría inteligente (ITS) advierte que los materiales educativos suelen reproducir modelos homogéneos, sin considerar diferencias culturales en la comprensión, la interpretación simbólica o las formas de aprendizaje [3, 4]. Esto impone el desafío de diseñar tecnologías educativas que fortalezcan las competencias interculturales y la apropiación territorial.

Los avances en inteligencia artificial generativa (IAGen) están transformando la forma en que se preserva, interpreta y transmite el patrimonio cultural. La IA es capaz de generar narrativas adaptativas, reconstrucciones históricas, experiencias multisensoriales y contenidos multimodales que fortalecen la interacción entre visitantes, comunidades y territorios [2, 4, 5]. Estos modelos no solo dinamizan la experiencia educativa, sino que

también promueven la inclusión, la sostenibilidad cultural y la accesibilidad [6], generando la posibilidad de crear tutores inteligentes capaces de ajustar sus narrativas y metáforas al territorio, lengua, cosmovisión y memoria histórica de cada comunidad [7].

2. DESARROLLO Y PROPUESTA

Se propone un modelo de ITS que se estructura mediante un flujo adaptativo que integra componentes pedagógicos, culturales y tecnológicos dentro de una arquitectura multiagente. El proceso comienza con la interacción del usuario a través de una interfaz conversacional, desde la cual se captura su intención y perfil contextual. Esta información es gestionada por la capa de orquestación adaptativa, que dirige cada solicitud hacia el agente especializado pertinente (pedagógico, narrativo, empático y de evaluación).

El agente pedagógico gestiona la secuencia didáctica, el agente cultural garantiza la pertinencia y la autenticidad del contenido, el agente narrativo construye relatos coherentes con la cosmovisión local, el agente empático modula la comunicación y el agente de evaluación monitorea. Estos agentes consultan una base de conocimiento biocultural, compuesta por narrativas orales, archivos históricos, relatos comunitarios de cada territorio, mediante un modelo de lenguaje (LLM). El sistema produce una salida personalizada, que retorna al usuario como tutoría o narrativa digital.

Como parte de la fase experimental, se implementó la arquitectura mediante flujos orquestados en n8n, integrando un modelo de lenguaje de OpenAI y cargando una base de conocimiento biocultural construida a partir de narrativas, testimonios y documentos relacionados con los pueblos indígenas del Ecuador. Para la interacción con los usuarios, se desarrolló una UI utilizando el framework Django, APIs REST y un webserver que permiten realizar consultas de forma accesible y estructurada. Esta integración posibilita que el sistema procese solicitudes, recupere información cultural relevante y genere respuestas adaptadas al contexto de manera eficiente desde cualquier navegador.

3. RESULTADOS

Se llevaron a cabo pruebas de usuario con el objetivo de evaluar el comportamiento del ITS durante la realización de consultas y la interacción con los distintos agentes del modelo. Estas pruebas permitieron observar la capacidad del sistema para interpretar adecuadamente las intenciones del usuario, recuperar información cultural relevante y generar respuestas adaptativas basadas en la base de conocimiento biocultural integrada. Los resultados obtenidos se presentan en la Figura 1.

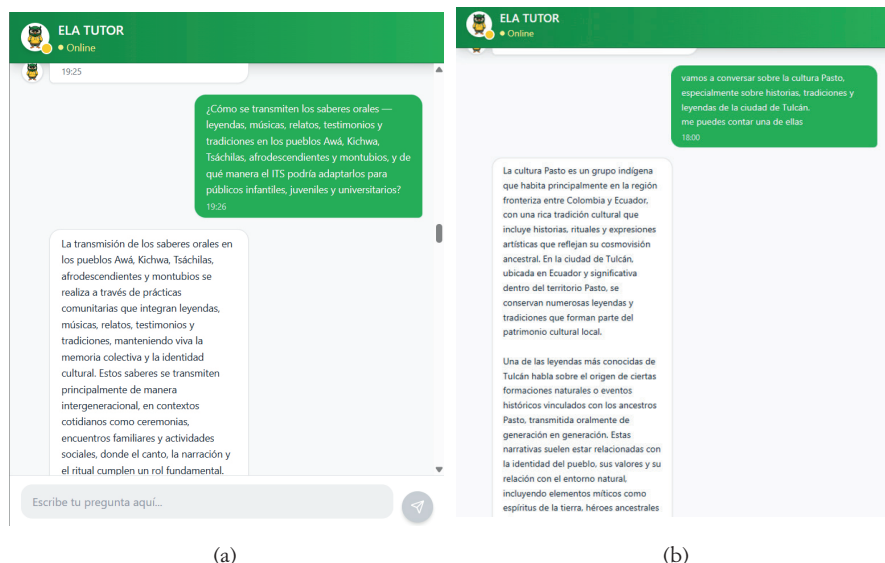


Figura 1. Interacción de ITS: (a) uso del agente pedagógico, (b) desarrollo de narrativa digital

4. CONCLUSIONES Y TRABAJOS FUTUROS

La integración de arquitecturas multiagente con modelos de IAGen representa una oportunidad significativa para transformar los procesos de enseñanza, conservación y transmisión de saberes ancestrales en contextos culturalmente diversos. El modelo propuesto demuestra que es posible articular tecnología con epistemologías y prácticas comunitarias, de manera que la IA no sustituya el conocimiento cultural, sino que funcione como un mediador respetuoso, contextual y adaptativo. El uso de una base de conocimiento biocultural y de un LLM supervisado por agentes pedagógicos y culturales garantiza que las respuestas generadas mantengan autenticidad y coherencia alineadas a los saberes ancestrales y las tradiciones orales de los pueblos. La propuesta explora el desarrollo de tecnologías educativas orientadas a la valorización del patrimonio biocultural y cosmovisión de los pueblos.

5. REFERENCIAS

1. Correia, J. E., et al. (2025). Biocultural geographies: Stewardship, Indigenous territories, and conservation in Ecuador's Amazon. *Annals of the American Association of Geographers*, 115(10), 2483-2501. <https://doi.org/10.1080/24694452.2025.2511942>

2. Liu, Y. (2025). Immersive digital storytelling to understand the new learning ecosystem: The revised concentric circles model from both teacher and students' perspective. *Smart Learning Environments*, 12(1), Artículo 39. <https://doi.org/10.1186/s40561-025-00398-x>
3. Mpondo, F. H., & Courtemanche, A. E. (2010). Cultural adaptation of pedagogical resources within intelligent tutorial systems. In V. Aleven & J. Kay (Eds.), *Intelligent Tutoring Systems* (pp. 340-342). Springer. https://doi.org/10.1007/978-3-642-13437-1_64
4. Blanchard, E. G., & Ogan, A. (2010). Infusing cultural awareness into intelligent tutoring systems for a globalized world. In R. Nkambou & J. Bourdeau (Eds.), *Advances in Intelligent Tutoring Systems* (pp. 485-505). Springer. https://doi.org/10.1007/978-3-642-14363-2_24
5. Sánchez-Martín, J.-M., Guillén-Peñafiel, R., & Hernández-Carretero, A.-M. (2025). Artificial intelligence in heritage tourism: Innovation, accessibility, and sustainability in the digital age. *Heritage*, 8(10), Artículo 428. <https://doi.org/10.3390/heritage8100428>
6. Filippone, A., Barbieri, U., Marsico, E., Bevilacqua, A., De Carlo, M. E., & Di Fuccio, R. (2025). Can 3D virtual worlds be used as intelligent tutoring systems to innovate teaching and learning methods? Future challenges and possible scenarios for metaverse and artificial intelligence in education. *Engineering Proceedings*, 87(1), Artículo 110. <https://doi.org/10.3390/engproc2025087110>
7. Kasemsarn, K., & Nickpour, F. (2025). Digital storytelling in cultural and heritage tourism: A review of social media integration and youth engagement frameworks. *Heritage*, 8(6), Artículo 200. <https://doi.org/10.3390/heritage8060200>

USO DE K-MEANS COMO HERRAMIENTA DE IA PARA LA PREVENCIÓN DE RIESGOS Y SÍNTOMAS RESPIRATORIOS EN POBLACIONES TRABAJADORAS VULNERABLES COLOMBIANAS

USE OF K-MEANS AS AN AI TOOL FOR THE PREVENTION OF RESPIRATORY RISKS AND SYMPTOMS IN VULNERABLE WORKING POPULATIONS IN COLOMBIA

Yenny Andrea ROZO SILVA¹, Leidy Isabel CALDERÓN SIERRA², Ana DELGADO-GARCÍA³,
Juan José AGÚN-GONZÁLEZ⁴ y Raúl AGUILAR ELENA⁵

¹ *Facultad de Ciencias Empresariales. Corporación Universitaria Minuto de Dios, Grupo de Investigación Ambientes Laborales Seguros, Saludables y Sostenibles (GIALSSS). Bogotá, Colombia*

✉ yrozo@uniminuto.edu

🌐 <https://orcid.org/0000-0002-3876-2372>

² *Facultad de Ciencias Empresariales. Corporación Universitaria Minuto de Dios, Grupo de Investigación Ambientes Laborales Seguros, Saludables y Sostenibles (GIALSSS). Bogotá, Colombia*

✉ icalderon@uniminuto.edu

🌐 <https://orcid.org/0000-0002-1056-901X>

³ *Grupo de Investigación BISITE, Universidad de Salamanca. Salamanca, España*

✉ a.delgado@usal.es

🌐 <https://orcid.org/0009-0003-7079-5507>

⁴ *Grupo de Investigación Prevención de Riesgos Laborales y Salud Laboral (GPRL), Cátedra PREVION. Universidad Internacional de Valencia (VIU). Valencia, España*

✉ jjagun@universidadviu.com

🌐 <https://orcid.org/0000-0002-2677-7198>

⁵ *Grupo de Investigación Prevención de Riesgos Laborales y Salud Laboral (GPRL), Cátedra PREVION. Universidad Internacional de Valencia (VIU). Valencia, España*

✉ raul.aguilar.e@professor.universidadviu.com

🌐 <https://orcid.org/0000-0001-8656-8155>

RESUMEN: El reciclaje informal en Colombia es una actividad ambientalmente necesaria, desarrollada en condiciones de elevada vulnerabilidad social y sanitaria, con exposición crónica

a riesgos biológicos y particulados y escaso acceso a vigilancia en salud laboral. En este contexto, la inteligencia artificial (IA) ofrece una oportunidad para identificar perfiles de riesgo ocultos en poblaciones sin datos etiquetados ni diagnósticos previos. Se realizó un estudio transversal en 179 recicladores informales de Bogotá, Tocancipá, Cajicá y Funza. Los síntomas respiratorios se evaluaron mediante el cuestionario ATS-DLD-78-A1 y la exposición y percepción del riesgo biológico mediante el cuestionario ERBIO. Los datos se procesaron en R y se analizaron mediante seis técnicas de *clustering* no supervisado (K-means, K-modes, jerárquico, DBSCAN, modelos de mezcla gaussiana y Fuzzy C-means), comparadas mediante métricas de validación interna. El algoritmo K-means mostró el mejor rendimiento global (índice de silueta = 0,639) e identificó tres clústeres: un grupo envejecido y altamente vulnerable con elevada prevalencia de síntomas respiratorios; un grupo joven con baja sintomatología; y un grupo intermedio, predominantemente femenino, con riesgo moderado. Otros métodos detectaron subgrupos extremos, pero con menor estabilidad e interpretabilidad. El *clustering* no supervisado, especialmente K-means, permite identificar perfiles diferenciados de riesgo respiratorio y orientar estrategias preventivas específicas. Integrada en la salud pública, la IA contribuye a la detección temprana del riesgo, la reducción de desigualdades y el avance hacia una prevención de precisión en salud laboral.

Palabras clave: inteligencia artificial; clustering no supervisado; síntomas respiratorios; trabajadores informales.

ABSTRACT: Informal recycling in Colombia is an environmentally necessary activity carried out under conditions of high social and health vulnerability, with chronic exposure to biological and particulate risks and limited access to occupational health surveillance. In this context, artificial intelligence (AI) offers an opportunity to identify hidden risk profiles in populations without labeled data or prior diagnoses. A cross-sectional study was conducted among 179 informal recyclers from Bogotá, Tocancipá, Cajicá, and Funza. Respiratory symptoms were assessed using the ATS-DLD-78-A1 questionnaire, and biological exposure and risk perception were evaluated using the ERBIO questionnaire. Data were processed in R and analyzed using six unsupervised clustering techniques (K-means, K-modes, hierarchical clustering, DBSCAN, Gaussian mixture models, and Fuzzy C-means), which were compared using internal validation metrics. The K-means algorithm showed the best overall performance (silhouette index = 0.639) and identified three clusters: an older, highly vulnerable group with a high prevalence of respiratory symptoms; a younger group with low symptomatology; and an intermediate, predominantly female group with moderate risk. Other methods detected extreme subgroups but demonstrated lower stability and interpretability. Unsupervised clustering, particularly K-means, enables the identification of differentiated respiratory risk profiles and the design of targeted preventive strategies. When integrated into public health, AI contributes to early risk detection, the reduction of inequalities, and progress toward precision prevention in occupational health.

Keywords: artificial intelligence; unsupervised clustering; respiratory symptoms; informal workers.

1. CONTEXTO Y JUSTIFICACIÓN

El reciclaje informal en Colombia constituye una actividad ambientalmente necesaria, pero desarrollada en condiciones de elevada vulnerabilidad social y sanitaria. Miles de personas dependen de esta labor sin protección social, expuestas de forma continuada a contaminantes y sin acceso efectivo a sistemas de vigilancia en salud [1,3]. Se trata de una población heterogénea y envejecida que refleja un patrón estructural de exclusión laboral. En este contexto, la inteligencia artificial ofrece un enfoque innovador para identificar perfiles de riesgo y patrones de enfermedad. El uso de algoritmos no supervisados, como K-means, permite analizar poblaciones diversas sin imponer categorías previas, aportando evidencia útil para la orientación de políticas públicas en salud laboral.

2. MATERIALES Y MÉTODOS

El estudio incluyó 179 recicladores informales de Bogotá, Tocancipá, Cajicá y Funza, seleccionados mediante muestreo no probabilístico por conveniencia. Se aplicaron el cuestionario ATS-DLD-78-A1 para síntomas respiratorios y el cuestionario ERBIO para exposición y percepción de riesgo biológico, ambos con adecuada consistencia interna ($\alpha > 0,70$). Las variables sociodemográficas y de salud se procesaron en R, descartando varianza nula y recodificando variables categóricas. Posteriormente se aplicaron seis técnicas de clustering (K-means, K-modes, jerárquico, DBSCAN, GMM y Fuzzy C-means), comparadas mediante métricas de validación interna (silueta, conectividad y Dunn).

3. RESULTADOS PRINCIPALES

El algoritmo K-means mostró el mejor rendimiento global (índice de silueta = 0,639), evidenciando una adecuada separación entre grupos. El número óptimo de clústeres fue $k = 3$, lo que permitió identificar tres perfiles poblacionales diferenciados. El Clúster 1 (23,8 %) correspondió a personas de mayor edad (media 64,8 años), con alta prevalencia de síntomas respiratorios y afiliación sanitaria subsidiada, reflejando un perfil de elevada vulnerabilidad. El Clúster 2 (17,9 %) agrupó a personas jóvenes (media 26,8 años), mayoritariamente hombres, con baja escolaridad y escasa sintomatología, representando un grupo laboralmente precario pero funcional. El Clúster 3 (20,4 %) incluyó adultos intermedios (media 46 años), con predominio femenino y presencia moderada de síntomas, asociado a una transición laboral marcada por riesgos de género. Otros métodos (K-modes, DBSCAN) identificaron subgrupos extremos, pero con menor estabilidad y consistencia interna. En conjunto, K-means proporcionó una segmentación clara, equilibrada y replicable, facilitando la interpretación preventiva. Los patrones sintomáticos por clúster se representaron mediante gráficos radar comparativos.

Tabla 1. Métricas de validación interna de los algoritmos de clustering aplicados

Method	Silhouette	Calinski–Harabasz	Davies–Bouldin	Dunn	Noise Ratio	Interpretation
K-Means	0.64	771.18	0.55	0.04	—	Compact, well-separated clusters
Hierarchical	0.64	768.34	0.55	0.09	—	Similar robust structure
GMM	0.29	12.94	3.48	0.15	—	Moderate overlap
K-Modes	—	29.22	2.32	0.03	—	Weak categorical structure
DBSCAN	0.08	10.41	3.94	0.01	0.06	Irregular, noise-tolerant
Fuzzy C-Means	0.04	7.08	5.38	0.01	—	Very diffuse boundaries

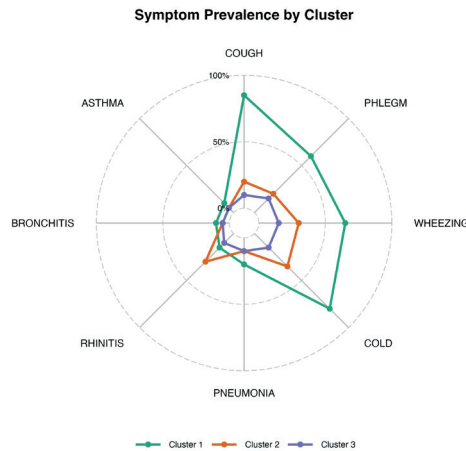


Figura 1. Distribución de síntomas respiratorios por clúster identificado mediante K-means.

4. DISCUSIÓN

Los resultados confirman el potencial del *clustering* como técnica de inteligencia artificial aplicada a la salud laboral. En particular, K-means permitió identificar perfiles de riesgo respiratorio no detectables mediante enfoques estadísticos tradicionales, resultando especialmente adecuado en contextos sin datos etiquetados ni diagnósticos previos, como las poblaciones trabajadoras informales.

La Figura 2 muestra un patrón consistente entre algoritmos: la presencia de al menos un clúster con alta prevalencia de síntomas respiratorios clave frente a otros grupos más

homogéneos y de menor riesgo [2,4]. Aunque métodos como GMM o Fuzzy C-means ofrecen transiciones más graduales y DBSCAN identifica casos extremos, K-means proporciona una segmentación clara, estable e interpretable.

Desde una perspectiva aplicada, K-means destaca por su equilibrio entre simplicidad computacional, robustez y facilidad de transferencia del conocimiento, facilitando el diseño de intervenciones preventivas dirigidas. Más allá del plano técnico, estos resultados evidencian el valor de la IA para visibilizar vulnerabilidades en poblaciones envejecidas, aportando una herramienta analítica con implicaciones éticas y de equidad en salud laboral y salud pública.

5. CONCLUSIONES

Este estudio confirma que la aplicación de inteligencia artificial mediante técnicas de *clustering* no supervisado permite identificar perfiles diferenciados de riesgo respiratorio en poblaciones trabajadoras informales, tradicionalmente invisibilizadas por los sistemas de vigilancia en salud. La segmentación obtenida facilita el diseño de intervenciones preventivas ajustadas a las características demográficas, laborales y sanitarias de cada grupo, superando el enfoque uniforme tradicional. El clúster más vulnerable requiere vigilancia respiratoria intensiva y medidas inmediatas de reducción de la exposición; el clúster de transición demanda acciones específicas con perspectiva de género; y el clúster resiliente se beneficia de estrategias de prevención temprana basadas en formación y control del riesgo. Integrada en políticas públicas y estrategias comunitarias, la inteligencia artificial se consolida como una herramienta eficaz para anticipar el daño, reducir desigualdades estructurales y avanzar hacia un modelo de prevención de precisión en salud laboral y salud pública [1,3].

6. REFERENCES

1. Binion, E., & Gutberlet, J. (2012). The effects of handling solid waste on the wellbeing of informal and organized recyclers: A review of the literature. *International Journal of Occupational and Environmental Health*, 18(1), 43-52. <https://doi.org/10.1179/1077352512Z.0000000001>
2. Ray, M. R., Roychoudhury, S., Mukherjee, G., Roy, S., & Lahiri, T. (2005). Respiratory and general health impairments of workers employed in a municipal solid waste disposal at an open landfill site in Delhi. *International Journal of Hygiene and Environmental Health*, 208(4), 255-262. <https://doi.org/10.1016/j.ijheh.2005.02.001>
3. Black, M., Karki, J., Lee, A. C. K., Makai, P., Baral, Y. R., Kritsotakis, E. I., & Turner, R. (2019). The health risks of informal waste workers in the Kathmandu Valley: A cross-sectional survey. *Public Health*, 166, 10-18. <https://doi.org/10.1016/j.puhe.2018.09.026>
4. Tlotleng, N., Kootbodien, T., Wilson, K., Made, F., Mathee, A., Ntlebi, V., & Naicker, N. (2019). Prevalence of respiratory health symptoms among landfill waste recyclers in the City of Johannesburg, South Africa. *International Journal of Environmental Research and Public Health*, 16(21), Artículo 4277. <https://doi.org/10.3390/ijerph16214277>

1.2.

PITCH DE PROYECTOS

ADAPTATION, ETHICAL TRACEABILITY AND REGULATORY COMPLIANCE IN CONVERSATIONAL ARTIFICIAL INTELLIGENCE AGENTS: AN INCLUSIVE RESEARCH FRAMEWORK FOR COGNITIVE AND SENSORY DIVERSITY

ADAPTACIÓN, TRAZABILIDAD ÉTICA Y CUMPLIMIENTO NORMATIVO EN AGENTES DE INTELIGENCIA ARTIFICIAL CONVERSACIONAL: UN MARCO DE INVESTIGACIÓN INCLUSIVO PARA LA DIVERSIDAD COGNITIVA Y SENSORIAL

Teresa CELESTINO-BUENTELLO¹ and Guillermo HERNÁNDEZ²

¹ *Air Institute. Salamanca, Castilla y León, Spain*

✉ mtcelestino@air-institute.com

🌐 <https://orcid.org/0000-0003-3984-7623>

² *Universidad de Salamanca. Salamanca, Castilla y León, Spain*

✉ guillehg@usal.es

🌐 <https://orcid.org/0000-0002-7481-5961>

ABSTRACT: This doctoral research proposes the development of an inclusive framework for conversational Artificial Intelligence that integrates functional adaptation, ethical traceability, and regulatory compliance. The research design addresses persistent accessibility gaps affecting users with speech, hearing, and cognitive disabilities through three empirical studies: the evaluation of speech recognition performance for disordered speech, the assessment of comprehension of synthetic versus natural speech among hearing-impaired users, and the analysis of cognitive adaptation in generative conversational agents. By combining quantitative and qualitative methods, the project establishes standardized metrics and ethical protocols aligned with the EU AI Act, WCAG 2.2, and the GDPR. By embedding accessibility and transparency throughout the AI lifecycle, this research advances evidence-based design principles for trustworthy, human-centered conversational systems, contributing to a more equitable and sustainable digital society.

Keywords: conversational AI; accessibility; trustworthy AI; AI act; cognitive diversity; sensory impairment; ethical compliance.

RESUMEN: Esta investigación doctoral propone el desarrollo de un marco inclusivo para la inteligencia artificial conversacional que integre la adaptación funcional, la trazabilidad ética y el cumplimiento normativo. El diseño de la investigación aborda las persistentes deficiencias de accesibilidad que afectan a los usuarios con discapacidades del habla, auditivas y cognitivas a través de tres estudios empíricos: la evaluación del rendimiento del reconocimiento del habla para el habla desordenada, la evaluación de la comprensión del habla sintética frente a la natural entre los usuarios con discapacidad auditiva y el análisis de la adaptación cognitiva en los agentes conversacionales generativos. Mediante la combinación de métodos cuantitativos y cualitativos, el proyecto establece métricas estandarizadas y protocolos éticos alineados con la Ley de IA de la UE, las WCAG 2.2 y el RGPD. Al incorporar la accesibilidad y la transparencia en todo el ciclo de vida de la IA, esta investigación promueve principios de diseño basados en la evidencia para sistemas conversacionales fiables y centrados en el ser humano, lo que contribuye a una sociedad digital más equitativa y sostenible.

Palabras clave: IA conversacional; accesibilidad; IA fiable; ley de IA; diversidad cognitiva; discapacidad sensorial; cumplimiento ético.

1. INTRODUCTION

Conversational Artificial Intelligence (AI) systems have become a dominant interface for mediating interactions between individuals and digital infrastructures in domains such as healthcare, education, and public administration. While recent advances in automatic speech recognition, natural language processing, and dialogue management have significantly improved overall system performance [1-3], most conversational agents continue to operate under a one-size-fits-all paradigm [4]. This design assumption privileges statistically normative communication patterns and, as a result, systematically excludes users whose speech, auditory perception, or cognitive processing diverge from these norms, thereby reinforcing existing digital inequalities. Empirical research has consistently shown that recognition accuracy for dysarthric or otherwise atypical speech remains substantially lower than for typical speakers [5-6], while synthetic or generated speech may introduce additional comprehension barriers for users with hearing or cognitive impairments [7].

In parallel, the European regulatory landscape has evolved toward a stricter conception of trustworthy AI. The Artificial Intelligence Act [8] identifies conversational systems deployed in sensitive domains as high-risk applications, explicitly requiring transparency, human oversight, and accessibility, in alignment with complementary frameworks such as WCAG 2.2 [9], EN 301 549, and the GDPR. These instruments emphasize that accessibility cannot be treated as a superficial interface feature but must be embedded into the design, evaluation, and governance of AI systems. Grounded in the social model of disability [10], which locates barriers in system design rather than in individual impairments [11-12], this doctoral research addresses the intersection of functional adaptation, ethical traceability, and legal compliance. It proposes an integrated methodological framework [13] to empirically evaluate and support inclusive conversational AI, addressing key gaps in accessibility

metrics, adaptive system behavior, and the responsible use of generative AI for cognitive and sensory diversity.

1.1. OBJECTIVES AND RESEARCH QUESTIONS

The doctoral project aims to develop and validate an inclusive conversational AI framework that integrates functional adaptation, ethical traceability, and compliance with European regulatory standards.

General Objective: To design, implement, and empirically evaluate an adaptive conversational AI system focused on spoken interaction, ensuring accessible, transparent, and lawful communication for users with cognitive and sensory diversity.

Specific Objectives: (1) To empirically evaluate the performance of automatic speech recognition (ASR) systems for speakers presenting phonological, motor, or articulatory impairments, establishing standardized metrics for accessibility and performance comparison. (2) To assess the comprehension and perceptual experience of synthetic versus natural speech among users with hearing impairment, identifying perceptual barriers and adaptation requirements for inclusive conversational interfaces. (3) To evaluate the cognitive adaptability of generative conversational AI systems (specifically GPT-based models) when supporting task execution and information comprehension for users with mild cognitive disabilities, ensuring compliance with accessibility and ethical standards.

2. METHODOLOGICAL FRAMEWORK

The research adopts a mixed-methods, multi-study design integrating quantitative and qualitative approaches. The empirical sequence comprises three interlinked studies addressing distinct gaps identified in the literature.

Study 1: Evaluation of automatic speech recognition performance in speakers with altered speech. This study evaluates the transcription accuracy of open-source ASR models, such as OpenAI's Whisper, using speech samples from participants with phonological, motor, or articulatory impairments (e.g., dysarthria, apraxia). Speech samples (both read and semi-spontaneous) are recorded and analyzed against a control group. Quantitative metrics include Character Error Rate (CER), substitution and omission rates, and systematic phoneme confusion, providing a detailed view of accuracy and error patterns in atypical speech. Statistical comparisons (t-tests, ANOVA) identify performance gaps and error typologies.

Study 2: Comprehension of synthetic versus natural speech in individuals with hearing impairment. This study evaluates auditory comprehension and listening effort when participants are exposed to natural versus text-to-speech (TTS) outputs. Individuals with mild-to-moderate hearing loss listen to paired messages identical in content but differing in source. Objective metrics (accuracy, response time) and subjective measures (perceived effort via adapted NASA-TLX, preference ratings) are analyzed using nonparametric tests where appropriate.

Study 3: Evaluation of cognitive adaptation in generative conversational agents. This study focuses on users with mild cognitive disabilities performing guided conversational tasks with generative conversational AI models (e.g., GPT-based systems), enabling comparison across architectures and adaptive behaviors. Linguistic complexity is controlled through distinct prompt configurations. Quantitative indicators are complemented by semi-structured interviews, with qualitative data analyzed through thematic analysis to ensure traceability and methodological rigor.

All studies include informed consent procedures, accessible materials, ethical review, and compliance with EU data protection regulations and research standards for vulnerable populations.

2.1. EXPECTED IMPACT AND ETHICAL IMPLICATIONS

This doctoral research generates scientific, social, and regulatory contributions through the integration of empirical evidence across three complementary studies. Scientifically, it advances the field of Inclusive and Trustworthy AI by proposing standardized metrics and methodological protocols for evaluating accessibility in conversational systems spanning speech recognition, speech synthesis, and cognitive interaction. These results support a more evidence-based understanding of how accessibility mechanisms influence usability, comprehension, trust, and fairness.

From a social perspective, the research addresses persistent exclusion in digital communication by improving access for users with speech, hearing, and cognitive disabilities. By supporting intelligible voice interfaces and cognitively accessible interactions, the project aligns with European priorities on digital inclusion and equal participation in education, healthcare, and public services.

Ethically and regulatorily, the project operationalizes key principles of the EU AI Act and the GDPR through ethical traceability, transparency, and human oversight. Privacy-by-design and human-in-the-loop supervision ensure accountability, data minimization, and risk mitigation in the use of generative AI. Overall, the research positions accessibility as a core requirement of trustworthy AI, offering a replicable framework for compliant, human-centered, and socially sustainable conversational systems.

3. ACKNOWLEDGMENTS

This research is part of the International Chair Project on Trustworthy Artificial Intelligence and the Demographic Challenge, within the framework of the National Artificial Intelligence Strategy (ENIA) and the European Recovery, Transformation and Resilience Plan. Reference: TSI-100933-2023-0001. This project is funded by the Secretary of State for Digitalization and Artificial Intelligence and by the European Union (Next Generation EU).

4. REFERENCES

1. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). *Robust speech recognition via large-scale weak supervision* (arXiv:2212.04356). arXiv. <https://doi.org/10.48550/arXiv.2212.04356>
2. Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Sengupta, S., Coates, A., & Ng, A. Y. (2014). *Deep Speech: Scaling up end-to-end speech recognition* (arXiv:1412.5567). arXiv. <https://doi.org/10.48550/arXiv.1412.5567>
3. Gulati, A., Qin, J., Chiu, C.-C., Parmar, N., Zhang, Y., Yu, J., Han, W., Wang, S., Zhang, Z., Pang, R., & Wu, Y. (2020). Conformer: Convolution-augmented transformer for speech recognition. In *Interspeech 2020 Proceedings* (pp. 5036-5040). International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2020-3015>
4. Vereschak, O., Alizadeh, Fatemeh., Bailly, G., & Caramiaux, B. (2024). Trust in AI-assisted decision making: Perspectives from those behind the system and those for whom the decision is made. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM. <https://doi.org/10.1145/3613904.3642018>
5. Calvo, I., Tropea, P., Viganò, M., Scialla, M., Cavalcante, A. B., Grajzer, M., Gilardone, M., & Corbo, M. (2021). Evaluation of an automatic speech recognition platform for dysarthric speech. *Folia Phoniatrica et Logopaedica*, 73(5), 427-435. <https://doi.org/10.1159/000511042>
6. Patel, T., Hutiri, W., Ding, A. Y., & Scharenborg, O. (2025). *How to evaluate automatic speech recognition: Comparing different performance and bias measures* [Manuscrito presentado para publicación].
7. Fu, B., Hadid, A., & Damer, N. (2025). Generative AI in the context of assistive technologies: Trends, limitations and future directions. *Image and Vision Computing*, 154, Artículo 105347. <https://doi.org/10.1016/j.imavis.2024.105347>
8. European Commission. (n.d.). *AI Act*. Retrieved June 23, 2025, from <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>
9. Kirkpatrick, A., O Connor, J., Campbell, A., & Cooper, M. (Eds.). (2025). *Web Content Accessibility Guidelines (WCAG) 2.1*. W3C Recommendation.
10. Sharma, T., Stangl, A., Zhang, L., Tseng, Y. Y., Xu, I., Findlater, L., Gurari, D., & Wang, Y. (2023). Disability-first design and creation of a dataset showing private visual information collected with people who are blind. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM. <https://doi.org/10.1145/3544548.3580922>
11. Bunton, K. (2008). Speech versus nonspeech: Different tasks, different neural organization. *Seminars in Speech and Language*, 29(4), 267-275. <https://doi.org/10.1055/s-0028-1103390>
12. Yorkston, K., & Beukelman, D. (1999). Staging interventions in progressive dysarthria. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, 9(4), 7-12. <https://doi.org/10.1044/nnsld9.4.7>
13. Schlosser, R. (2009). Social validation of interventions in augmentative and alternative communication. *Augmentative and Alternative Communication*, 15(4), 234-247. <https://doi.org/10.1080/07434619912331278775>

EDIFICIOS INTELIGENTES EN ENTORNOS RURALES: INTEGRACIÓN DE IOT E INTELIGENCIA ARTIFICIAL EN LA PROVINCIA DE VALLADOLID

SMART BUILDING IN RURAL AREAS: INTEGRATION OF IOT AND ARTIFICIAL INTELLIGENCE IN THE PROVINCE OF VALLADOLID

David ESCUDERO-MANCEBO¹ y Jezabel GÓMEZ²

¹ *Universidad de Valladolid. Valladolid, Castilla y León, España*

✉ escuderomancebo.david@uva.es

 <https://orcid.org/0000-0003-0849-8803>

² *Universidad de Valladolid. Valladolid, Castilla y León, España*

 <https://orcid.org/1111-2222-3333-4444>

RESUMEN: El proyecto de despliegue de edificios inteligentes impulsado por la Diputación de Valladolid tiene como objetivo contribuir al desarrollo del entorno rural mediante la digitalización. El Centro IA de la Universidad de Valladolid se ha incorporado al mismo para explorar las posibilidades del uso de tecnologías de inteligencia artificial en dicho fin. Se presenta el recorrido del proyecto, desde la instalación de sensores IoT en edificios municipales hasta la creación de un *dataset* multivariante con datos energéticos, ambientales y de ocupación. Se describen los procesos de ingesta y almacenamiento de la información en bases de datos de series temporales, así como las primeras aplicaciones de minería de datos y técnicas de *machine learning* para la detección de patrones y optimización de consumos. El trabajo concluye con una reflexión sobre el valor estratégico de la infraestructura IoT rural y su papel en la cohesión territorial y el reto demográfico mediante la IA.

Palabras clave: IoT; inteligencia artificial; territorios inteligentes.

ABSTRACT: The smart building deployment project promoted by the Valladolid Provincial Council aims to contribute to the development of rural areas through digitalisation. The University of Valladolid's AI Centre has joined the project to explore the possibilities of using artificial intelligence technologies for this purpose. The project's progress is presented, from the installation of IoT sensors in municipal buildings to the creation of a multivariate

dataset with energy, environmental and occupancy data. The processes of ingesting and storing information in time series databases are described, as well as the first applications of data mining and machine learning techniques for pattern detection and consumption optimisation. The paper concludes with a reflection on the strategic value of rural IoT infrastructure and its role in territorial cohesion and the demographic challenge through AI.

Keywords: IoT; artificial intelligence; smart territories.

1. INTRODUCCIÓN

El desarrollo de edificios inteligentes en el medio rural surge como respuesta a una combinación de retos estructurales que afectan a estos territorios, entre los que destacan la despoblación, la limitación de infraestructuras técnicas y el creciente impacto de los costes energéticos en la gestión municipal [1]. La digitalización se plantea como una oportunidad estratégica para reforzar la cohesión territorial, mejorar la eficiencia de los servicios públicos y avanzar hacia modelos de gestión más sostenibles [2, 3]. En este contexto se enmarca la iniciativa impulsada por la Diputación de Valladolid cuyo objetivo es dotar a los edificios municipales de una infraestructura IoT que permita monitorizar el consumo energético, las condiciones ambientales y los usos reales de los espacios¹. Esta base tecnológica no solo posibilita una mejor toma de decisiones, sino que abre la puerta a explorar progresivamente aplicaciones de inteligencia artificial, sentando las bases de una estrategia digital rural orientada a la modernización, la eficiencia y la mejora de la calidad de vida en los municipios.

2. RECURSOS Y MÉTODOS

La Figura 1 muestra el despliegue tecnológico en los edificios municipales, incluyendo sensores de consumo energético, calidad del aire (CO₂), temperatura, humedad y presencia, entre otros, con el fin de capturar información continua sobre el funcionamiento real de las instalaciones. Estos dispositivos se integran mediante pasarelas de comunicación que garantizan la conectividad y el envío de datos a plataformas centralizadas, empleando tecnologías adaptadas a las condiciones del medio rural. Los datos se almacenan en una base de datos SQL a la cual se accede para ofrecer servicios de visualización de información mediante un API REST.

En la Tabla 2 se describen más de 700 sensores. Se trata de un conjunto integral de sensores y dispositivos IoT orientados a la monitorización y la gestión inteligente de edificios y espacios urbanos, que permiten medir y optimizar consumos energéticos (electricidad, energía térmica, agua, gas y gasoil), controlar condiciones ambientales y de confort (temperatura, calidad del aire, iluminación, ruido y meteorología), gestionar infraestructuras y servicios (climatización, iluminación interior, aparcamientos, residuos y

1 <https://www.diputaciondevalladolid.es/web/soporte/edificios-inteligentes>

estaciones meteorológicas) y mejorar la seguridad y el uso de los espacios mediante el control de accesos, la detección de inundaciones y el análisis de afluencia y movimiento de personas.

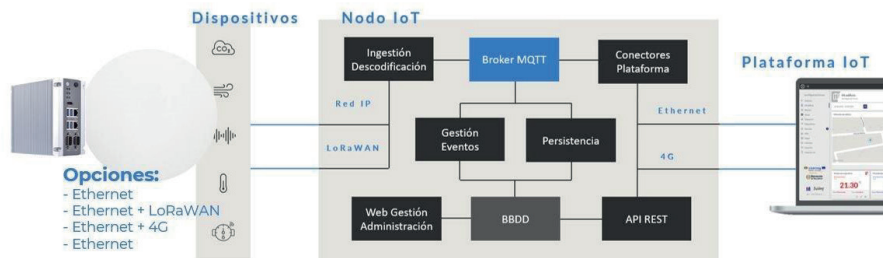


Figura 1. Arquitectura desplegada para la captura, el almacenamiento y la visualización de la información.

[illegible]

Tab. 1. Tipología de los sensores y su distribución geográfica

3. POTENCIALES APLICACIONES DE IA Y MACHINE LEARNING

En la aplicación de técnicas de inteligencia artificial y de *machine learning* en este contexto, el principal desafío identificado es la falta de datos etiquetados [3], una limitación habitual en entornos municipales donde, pese a contar ya con sensorica desplegada y datos disponibles, la aplicación avanzada de IA aún no se ha consolidado. En este escenario, las aproximaciones no supervisadas, basadas en parametrización y técnicas de *clustering* [4], se perfilan como una alternativa viable a corto plazo. Una primera exploración de los resultados evidencia diferencias significativas en las series temporales de la misma magnitud asociadas a diferentes espacios (tanto en el mismo edificio como entre edificios). El uso de referencias parametrizadas de las series temporales y algoritmos clásicos de detección de anomalías [5] se perfila como una técnica a ser explorada.

Asimismo, los modelos de predicción de momentos críticos de mantenimiento [6] representan un problema más acotado y, por tanto, más sencillo de abordar en fases iniciales, lo que explica su alta aceptación por parte de los municipios [7] como primer paso hacia una adopción progresiva de soluciones de inteligencia artificial. Casos de uso tales como el envío de avisos cuando haya que reponer el gasoil de la caldera han demostrado ser bien percibidos por la población y los agentes beneficiarios de las tecnologías.

Los datos han sido exportados a la base de datos *InfluxDB* para usar *Grafana* como *dashboard*. En paralelo se han creado archivos de características con las series temporales. El volumen de información es elevado (se monitorizan 6 muestras por hora), lo cual permitiría explorar el uso de modelos preentrenados sobre series temporales multivariantes [3, 4] como alternativa al uso de parametrizaciones (histogramas, LPC y/o FFT, por ejemplo). La disponibilidad de métricas objetivas derivadas de la sensórica abre además la puerta a evaluar de manera sistemática la calidad y la eficacia de las intervenciones realizadas, en línea con enfoques de evaluación cuantitativa [8], permitiendo no solo justificar decisiones de gestión, sino también sentar las bases para procesos de optimización continua asistidos por inteligencia artificial

4. CONCLUSIONES

La digitalización de los entornos rurales se presenta como una opción estratégica clave para su desarrollo, al tiempo que se convierte en un elemento necesario para impulsar la modernización municipal. En este contexto, el IoT aplicado a edificios inteligentes permite recopilar grandes volúmenes de datos a través de la sensórica, cuyo valor se ve reforzado por la inteligencia artificial, llamada a desempeñar un papel fundamental en su análisis y explotación. Esta combinación tecnológica contribuye de manera directa al reto demográfico, ya que posibilita el envío de alertas que evitan la presencia in situ, facilita el control remoto asistido de las instalaciones y genera importantes ahorros en la gestión y el mantenimiento. Tener en cuenta consideraciones de seguridad [9] y de calidad del dato [10] debe traducirse en una percepción positiva tanto por parte de las autoridades locales como de los vecinos, que ven en estas soluciones una mejora tangible de los servicios y de la calidad de vida en el entorno rural.

5. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Asesoramiento en el uso de sistemas inteligentes para la gobernanza del dato en entornos provinciales: caso de uso en la provincia de Valladolid (SIGODEP-VA). Referencia: FECYT clave FCT-24-21134.

6. REFERENCIAS

1. Mosannenzadeh, F., & Vettorato, D. (2014). Defining smart city. A conceptual framework based on keyword analysis. *TeMA - Journal of Land Use, Mobility and Environment*, 7(3), 257-276. <https://doi.org/10.6092/1970-9870/2523>
2. Gómez-Valenzuela, V., & Holl, A. (2025). Rural depopulation in the context of 4.0 technologies: Opportunities for sustainability and innovation policies. *Journal of Planning Literature*, 40(3), 407-424. <https://doi.org/10.1177/08854122241276015>
3. Liu, Y., Zhou, Y., Yang, K., & Wang, X. (2023). Unsupervised deep learning for IoT time series. *IEEE Internet of Things Journal*, 10(16), 14285-14306. <https://doi.org/10.1109/JIOT.2023.3243391>
4. Dridi, J., Amayri, M., & Bouguila, N. (2024). Unsupervised clustering-based domain adaptation for estimating occupancy and recognizing activities in smart buildings. *Journal of Building Engineering*, 85, Artículo 108741. <https://doi.org/10.1016/j.jobbe.2024.108741>
5. Cook, A. A., Mısırlı, G., & Fan, Z. (2019). Anomaly detection for IoT time-series data: A survey. *IEEE Internet of Things Journal*, 7(7), 6481-6494. <https://doi.org/10.1109/JIOT.2019.2958185>
6. Sater, R. A., & Hamza, A. B. (2021). A federated learning approach to anomaly detection in smart buildings. *ACM Transactions on Internet of Things*, 2(4), 1-23. <https://doi.org/10.1145/3467981>
7. Dissem, M., Amayri, M., & Bouguila, N. (2024). Neural architecture search for anomaly detection in time-series data of smart buildings: A reinforcement learning approach for optimal autoencoder design. *IEEE Internet of Things Journal*, 11(10), 18059-18073. <https://doi.org/10.1109/JIOT.2024.3360882>
8. Morano, P., Guarini, M. R., Tajani, F., & Anelli, D. (2020). Sustainable redevelopment: The cost-revenue analysis to support the urban planning decisions. In O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, & C. M. Torre (Eds.), *Computational Science and Its Applications – ICCSA 2020* (pp. 968-1080). Springer. https://doi.org/10.1007/978-3-030-58808-3_69
9. Alonso-Núñez, J., López Martínez, D., & Llanos Ferraris, D. R. (2025). *Arquitectura segura para la trazabilidad basada en IoT y blockchain* [Manuscrito presentado para publicación].
10. Pezzutto, S., Bottino-Leone, D., & Wilczynski, E. J. (2025). Enhancing sustainability through quality controlled energy data: The Horizon 2020 EnerMaps Project. *Sustainability*, 17(17), Artículo 7684. <https://doi.org/10.3390/su17177684>

GRANDES MODELOS DE LENGUAJE COMO MEDIADOR LINGÜÍSTICO EN LA ATENCIÓN MÉDICA

LARGE LANGUAGE MODELS AS LANGUAGE MEDIATORS IN HEALTHCARE

Thalia RUIZ DESDÍN

Universidad de Salamanca. Salamanca, Castilla y León, España

✉ truiz@usal.es

🌐 <https://orcid.org/0009-0001-6673-3106>

RESUMEN: Los grandes modelos de lenguaje (LLM) están transformando el sector de la salud no solo en el diagnóstico y la predicción de enfermedades, sino también en la educación y la comunicación con los pacientes. La complejidad del lenguaje médico, sumada a diferencias en el nivel de alfabetización en salud y a las barreras idiomáticas, dificulta la comprensión de diagnósticos y tratamientos por parte de la población general. En este contexto, los LLM emergen como herramientas con potencial para actuar como mediadores lingüísticos entre médicos y pacientes.

Dado su ampliación a otras prestaciones de la atención sanitaria, este trabajo se propone analizar cómo se usan los LLM para facilitar la comunicación con los pacientes. Con esta finalidad, se revisaron estudios recientes centrados en aplicaciones clínicas y educativas de modelos como ChatGPT, principalmente en las especialidades de radiología, ortopedia y oncología.

Los resultados muestran que los LLM presentan una alta capacidad para simplificar información médica compleja, generar explicaciones comprensibles y apoyar la toma de decisiones informadas por parte del paciente. No obstante, también se identifican limitaciones relacionadas con la precisión diagnóstica, el rendimiento en idiomas distintos al inglés y consideraciones éticas y de privacidad.

Sin embargo, a pesar de estas limitaciones, se concluye que los LLM poseen un alto potencial como herramientas complementarias en la atención médica, especialmente en la comunicación interlingüística entre médico y paciente, evidenciando la necesidad de realizar estudios experimentales en un futuro para comprobar su factibilidad.

Palabras clave: grandes modelos del lenguaje; mediador lingüístico; consultas médicas.

ABSTRACT: Large language models (LLMs) are transforming the healthcare sector, not only in the diagnosis and prediction of diseases, but also in patient education and communication.

The complexity of medical language, coupled with differences in health literacy levels and language barriers, makes it difficult for the general population to understand diagnoses and treatments. In this context, LLMs are emerging as tools with the potential to act as linguistic mediators between doctors and patients.

Given their expansion into other healthcare services, this paper aims to analyse how LLMs are used to facilitate communication with patients. To this end, recent studies focusing on clinical and educational applications of models such as ChatGPT were reviewed, mainly in the specialities of radiology, orthopaedics and oncology.

The results show that LLMs have a high capacity to simplify complex medical information, generate understandable explanations, and support informed decision-making by patients. However, limitations related to diagnostic accuracy, performance in languages other than English, and ethical and privacy considerations were also identified.

Despite these limitations, it is concluded that LLMs have high potential as complementary tools in healthcare, especially in communication.

The results show that LLMs have a high capacity to simplify complex medical information, generate understandable explanations, and support informed decision-making by patients. However, limitations related to diagnostic accuracy, performance in languages other than English, and ethical and privacy considerations are also identified.

Despite these limitations, it is concluded that LLMs have high potential as complementary tools in healthcare, especially in interlinguistic communication between doctors and patients, highlighting the need for future experimental studies to test their feasibility.

Keywords: Major language models; language mediator; medical consultations.

1. INTRODUCCIÓN

Hasta la fecha se aprecia cómo la incorporación de los grandes modelos de lenguaje en el sector de la salud genera un impacto significativo, tanto en los procedimientos médicos como en la interacción entre sus profesionales y los pacientes. En un principio estas tecnologías se implementaron mediante modelos predictivos orientados a reducir tiempos de atención y mejorar la precisión de los diagnósticos. Sin embargo, su aplicación se ha ampliado progresivamente hacia la educación sanitaria y la comunicación clínica.

Para los pacientes sin formación en ciencias de la salud, la comprensión de diagnósticos, informes médicos y tratamientos complejos resulta frecuentemente limitada debido al uso de terminología técnica, procesos biológicos complejos y explicaciones poco adaptadas a su nivel de conocimiento. Esta problemática se intensifica en contextos multilingües, como en el caso de pacientes migrantes o turistas. En este escenario, los LLM pueden desempeñar un rol clave como mediadores lingüísticos, facilitando una comunicación más clara, accesible y personalizada.

2. METODOLOGÍA

En este trabajo se realizó una revisión bibliográfica de estudios recientes que analizan el uso de LLM en el ámbito sanitario, con énfasis en su aplicación en la comunicación médico-paciente. Se incluyeron investigaciones centradas en modelos de uso general entre la población, como ChatGPT (OpenAI), Gemini (Google), Copilot (Microsoft) y Claude (Anthropic). También se revisaron estudios que describen modelos especializados en medicina, aunque el análisis principal se centró en plataformas de uso general, debido a su mayor accesibilidad y adopción por parte de la ciudadanía. Los trabajos seleccionados abarcan distintas especialidades médicas y evalúan aspectos como precisión, usabilidad, claridad del lenguaje y fiabilidad de la información proporcionada.

3. RESULTADOS Y DISCUSIÓN

La literatura revisada evidencia que hasta la fecha los LLM se han utilizado para interpretar información de pruebas como la radiológica. Kuckelman *et al.* [1] resaltaron la precisión que ha mostrado Chat GPT en la realización de informes de radiología musculoesquelética. En la especialidad de ortopedia, Xue *et al.* [2] descubrieron que también es de utilidad en la realización de consultas remotas a través de un estudio que involucró a médicos certificados en esta especialidad como grupo de control y a Chat GPT como experimental, obteniendo diagnósticos similares.

Afin con la línea de investigación del estudio también se encontraron los resultados de An *et al.* [3], donde se introdujo un modelo educativo basado en LLM dirigido a facilitar la comprensión del estado de salud y tratamiento por parte de los pacientes oncológicos. De esta forma educa al enfermo en aspectos relacionados con el examen médico, la atención sanitaria, la radioterapia y la quimioterapia. Este modelo está enfocado además en preparar al médico para las diferentes preguntas que puede realizarle el paciente o sus familiares durante la consulta, con vistas a aclarar cada una de sus dudas de una forma asequible a su nivel de conocimiento.

A pesar de los diferentes beneficios en el uso de los LLM, los estudios señalan diversas limitaciones, entre ellas, posibles sesgos en la información, menor rendimiento en idiomas distintos al inglés, tendencia a reforzar patrones probabilísticos en lugar de razonamiento clínico y riesgos asociados a la simplificación excesiva del lenguaje médico. A ello se suman preocupaciones éticas y de privacidad del paciente, que deben ser consideradas en cualquier implementación clínica.

4. CONCLUSIONES

Los estudios analizados confirman el potencial de los LLM como herramientas de apoyo en la educación del paciente y la optimización de la atención sanitaria. Sin embargo,

se observa una ausencia de investigaciones que evalúen de forma específica su uso como mediadores lingüísticos directos en la consulta médica, especialmente en contextos multilingües y con distintos niveles de alfabetización en salud.

Esta revisión justifica la pertinencia de futuras investigaciones experimentales destinadas a evaluar la efectividad y la fiabilidad de los LLM en estos escenarios, así como a analizar de manera sistemática los sesgos y los riesgos asociados a su uso en entornos clínicos reales.

5. AGRADECIMIENTOS

Agradezco a mis tutores M. Angélica González Arrieta y a Guillermo Hernández González por ayudarme y asesorarme en esta aventura que recién comienza.

6. REFERENCIAS

1. Kuckelman, I. J., Wetley, K., Yi, P. H., & Ross, A. B. (2024). Translating musculoskeletal radiology reports into patient-friendly summaries using ChatGPT-4. *Skeletal Radiology*, 53(8), 1621-1624. <https://doi.org/10.1007/s00256-024-04599-2>
2. Xue, Z., Zhang, Y., Gan, W., Wang, H., She, G., & Zheng, X. (2024). Quality and dependability of ChatGPT and Dingxiangyuan forums for remote orthopedic consultations: Comparative analysis. *Journal of Medical Internet Research*, 26, Artículo e50882. <https://doi.org/10.2196/50882>
3. An, Y., Fang, Q., & Wang, L. (2024). Enhancing patient education in cancer care: Intelligent cancer patient education model for effective communication. *Computers in Biology and Medicine*, 169, Artículo 107874. <https://doi.org/10.1016/j.compbiomed.2023.107874>

DESAFÍOS DE LA IA GENERATIVA EN DOMINIOS DE ALTO RIESGO: REQUISITOS PARA LA SIMPLIFICACIÓN JUDICIAL FIABLE

CHALLENGES OF GENERATIVE AI IN HIGH-RISK DOMAINS: REQUIREMENTS FOR RELIABLE JUDICIAL SIMPLIFICATION

Ana-Belén GIL-GONZÁLEZ¹, Silvia DOMÍNGUEZ CASTRO², Clara RODRÍGUEZ MARCOS³,
Alberto JIMÉNEZ SÁNCHEZ⁴, Javier MANSO CONDE⁵ y Jorge Enrique GONZÁLEZ GONZALO⁶

¹ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ abg@usal.es

🌐 <https://orcid.org/0000-0001-7235-6151>

² *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ sdominguez@air-institute.com

🌐 <https://orcid.org/0009-0006-5290-7394>

³ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ crodriguez@air-institute.com

🌐 <https://orcid.org/0009-0005-2713-9269>

⁴ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ ajimenez@air-institute.com

🌐 <https://orcid.org/0009-0004-4183-0008>

⁵ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jmanso@air-institute.com

🌐 <https://orcid.org/0009-0005-5187-9566>

⁶ *Universidad de Salamanca. Salamanca, Castilla y León, España*

✉ jgonzalez@air-institute.com

🌐 <https://orcid.org/0009-0004-4053-3503>

RESUMEN: El acceso a la justicia requiere no solo la posibilidad de acudir a un tribunal, sino la capacidad de comprender sus decisiones. En España, la Ley 6/2022 establece la accesibilidad cognitiva como un derecho, promoviendo el uso de la lectura fácil (LF). Sin embargo, la saturación del sistema judicial hace inviable la adaptación manual masiva de sentencias. La inteligencia artificial generativa (GenIA) surge como una solución potencial,

pero su aplicación en justicia es clasificada como de «alto riesgo» por el Reglamento de IA de la UE debido a problemas como las alucinaciones y la falta de explicabilidad. Este trabajo analiza dichos desafíos técnicos y jurídicos, proponiendo una arquitectura de fiabilidad basada en tres requisitos innegociables: anclaje factual con trazabilidad documental, supervisión humana obligatoria y segregación funcional de roles. El objetivo es definir un estándar que permita automatizar la simplificación del lenguaje jurídico sin comprometer el rigor legal ni las garantías procesales.

Palabras clave: inteligencia artificial generativa; lectura fácil; acceso a la justicia.

ABSTRACT: Access to justice requires not only the possibility of going to court, but also the ability to understand its decisions. In Spain, Law 6/2022 establishes cognitive accessibility as a right, promoting the use of Easy Reading (LF). However, the saturation of the judicial system makes the manual adaptation of judgments unfeasible. Generative Artificial Intelligence (GenAI) emerges as a potential solution, but its application in justice is classified as «High Risk» by the EU AI Regulation due to problems such as hallucinations and lack of explainability. This paper analyses these technical and legal challenges, proposing a reliability architecture based on three non-negotiable requirements: factual anchoring with documentary traceability, mandatory human supervision and functional segregation of roles. The aim is to define a standard that allows the simplification of legal language to be automated without compromising legal rigour or procedural guarantees.

Keywords: generative artificial intelligence; easy reading; access to justice.

1 INTRODUCCIÓN

El acceso a la justicia no consiste solo en la posibilidad de acudir a un tribunal, sino en la capacidad de comprender las decisiones que allí se toman. Sin embargo, el lenguaje jurídico actual suele ser oscuro, excesivamente técnico y difícil de entender para el ciudadano medio. Esta complejidad crea una barrera de comunicación que afecta a los colectivos más vulnerables, limitando su autonomía para entender sus propios derechos y obligaciones [1, 18].

En España, la legislación ha dado un paso decisivo para resolver este problema. La aprobación de la Ley 6/2022 [12] ha convertido la accesibilidad cognitiva en un derecho exigible, obligando a las Administraciones públicas a adaptar sus comunicaciones para que sean comprensibles. Esto incluye el uso de la lectura fácil (LF), una técnica de redacción y diseño regulada por la norma UNE 153101:2018 [2]. No obstante, la aplicación de esta normativa se enfrenta a un obstáculo: el sistema judicial está saturado y la adaptación manual de cada sentencia es una tarea inviable debido al enorme volumen de documentos que se generan a diario [6].

Ante este escenario, la inteligencia artificial generativa (GenIA) se presenta como una herramienta capaz de realizar esta tarea. Sin embargo, el uso de modelos de lenguaje estándar (como los utilizados en ChatGPT) conlleva riesgos en el ámbito legal, debido a su tendencia a inventar información (las llamadas «alucinaciones») y a la falta de transparencia sobre cómo

toman decisiones [3]. Por este motivo, el reciente Reglamento de Inteligencia Artificial de la Unión Europea [11] clasifica el uso de la inteligencia artificial (IA) en la justicia como un sistema de «alto riesgo», exigiendo garantías estrictas de supervisión y fiabilidad.

El objetivo de este trabajo es analizar estos desafíos y definir qué requisitos técnicos debe cumplir un sistema de IA para automatizar la LF sin entrar en conflicto con el rigor legal.

2. EL IMPERATIVO NORMATIVO: DE LA RECOMENDACIÓN A LA OBLIGACIÓN

Durante décadas, la simplificación del lenguaje administrativo en España se mantuvo en el terreno de las recomendaciones de estilo o los códigos de buenas prácticas [1]. Sin embargo, el marco jurídico ha sufrido una transformación radical donde el punto de inflexión fue la Ley 6/2022, de 31 de marzo, que modifica el Texto Refundido de la Ley General de derechos de las personas con discapacidad [12]. Esta norma es trascendental porque define y regula, por primera vez, la accesibilidad cognitiva dentro del ordenamiento jurídico español, estableciendo que la Administración de justicia debe garantizar que sus procesos sean comprensibles para todas las personas, independientemente de sus capacidades intelectuales [4, 24]. Ya estamos ante un derecho subjetivo exigible ante los tribunales, cuyo incumplimiento puede derivar en la nulidad de actuaciones por indefensión.

La accesibilidad se ha tecnificado a través de estándares como la norma UNE 153101:2018 [2], la cual aclara una confusión habitual en la práctica forense: la LF no es un mero resumen ni una versión «infantilizada» del texto original [5]. Se trata de una técnica de redacción y publicación que impone reglas estrictas, tales como la limitación de la longitud de las oraciones, el uso de glosarios laterales para términos complejos y la eliminación de metáforas o lenguaje figurado que puedan distorsionar el mensaje [2, 19]. La norma subraya que la adaptación no es solo lingüística, sino también visual, exigiendo tipografías específicas, márgenes amplios y una jerarquía de la información que facilite el rastreo visual, elementos que la Ley 6/2022 ha convertido en condiciones básicas de accesibilidad [12, 21].

En este contexto, la adaptación documental converge con la disciplina académica del *Legal Design* (Diseño Legal). Este enfoque, teorizado por investigadoras como Margaret Hagan [10], señala que la justicia debe ser diseñada «centrada en el usuario» para ser efectiva. Aplicado a las sentencias, el *Legal Design* implica que la simplificación no se limita a reescribir párrafos, sino a reestructurar la información para hacerla navegable y reducir la carga cognitiva [9]. Estudios recientes demuestran que la combinación de un lenguaje claro y de un diseño visual mejora la capacidad de las personas para tomar decisiones legales informadas [13]. Por consiguiente, cualquier sistema de IA no puede limitarse a ser un «traductor» de palabras; debe operar con una interoperabilidad técnica y procedimental que le permita actuar como un diseñador de información, organizando visualmente el contenido para garantizar la comprensión [20, 23].

3. RIESGOS DE LA IA GENERATIVA EN DOMINIOS DE ALTO RIESGO

La integración de grandes modelos de lenguaje (LLM) en la Administración Pública no puede realizarse de manera acrítica. El ámbito jurídico exige un estándar de precisión factual y seguridad normativa que los modelos generativos actuales no pueden garantizar [25]. Esta incompatibilidad ha motivado que el legislador europeo, en el reciente Reglamento (UE) 2024/1689 [17], clasifique explícitamente en su Anexo III los sistemas de IA destinados a la Administración de Justicia como sistemas de «alto riesgo» [16]. Esta clasificación no es arbitraria, sino que responde a tres amenazas sistémicas que comprometen las garantías del debido proceso.

En primer lugar, la naturaleza de los LLM da lugar al fenómeno de las «alucinaciones». Tal como advierten Bender *et al.* [3], estos modelos no acceden a una base de datos de verdad, sino que operan prediciendo la siguiente palabra más probable basándose en patrones estadísticos, sin conexión con la realidad. En el dominio legal, estudios recientes realizados por la Universidad de Stanford han demostrado que los modelos de propósito general (como GPT-4) inventan jurisprudencia o citan normativas inexistentes en un porcentaje alarmante de las consultas jurídicas complejas (entre el 15 % y el 25 %), produciendo una veracidad aparente, pero factualmente falsa [8]. En una sentencia de LF, una alucinación que altere un hecho probado o la duración de una pena no es un error informático, sino una vulneración directa de la tutela judicial efectiva.

En segundo lugar, la opacidad algorítmica o el problema de la «Caja Negra» contraviene el principio de motivación de las resoluciones judiciales. El artículo 24 de la Constitución española y la normativa europea exigen que cualquier decisión administrativa sea explicable y trazable [7]. Sin embargo, el funcionamiento interno de las redes neuronales profundas es ininteligible para el operador humano, lo que impide auditar por qué el sistema decidió simplificar un término de una manera y no de otra. El Grupo de Expertos de Alto Nivel de la UE [11] ha señalado que esta falta de explicabilidad es inaceptable en sectores donde se dirimen derechos fundamentales, exigiendo mecanismos de supervisión humana obligatorios.

Finalmente, existe el riesgo de la sobresimplificación, debido a que el léxico jurídico posee una carga semántica precisa donde la eliminación de matices altera la sustancia del mensaje. Investigaciones sobre *Legal Tech* advierten que los modelos entrenados para maximizar la legibilidad tienden a eliminar cláusulas subordinadas o términos técnicos que son necesarios en una sentencia [22]. Una IA que priorice la facilidad de lectura por encima del rigor legal podría producir textos muy accesibles, pero jurídicamente vacíos o incorrectos, despojando al ciudadano de la información necesaria para su defensa y creando una falsa sensación de comprensión [15].

4. REQUISITOS PARA UNA IA DE SIMPLIFICACIÓN FIABLE

Con el objetivo de que la automatización de la LF sea compatible con el Estado de derecho, cualquier arquitectura propuesta debe satisfacer, como mínimo, tres condiciones de fiabilidad que mitiguen los riesgos inherentes a la GenIA.

4.1. REQUISITO 1: ANCLAJE FACTUAL Y TRAZABILIDAD DOCUMENTAL

El principio de veracidad judicial exige que la IA abandone su comportamiento creativo. El modelo no debe generar texto basándose en su entrenamiento previo (que puede contener sesgos), sino exclusivamente en la información contenida en la sentencia original. Técnicamente, esto requiere mecanismos de recuperación de información que obliguen al algoritmo a citar o referenciar el fragmento exacto del documento fuente que justifica cada frase simplificada [14]. Si el sistema no puede trazar el origen de una afirmación hasta el texto original, no debe generarla. Solo mediante esta trazabilidad documental es posible realizar la auditoría *ex post* que exige el Reglamento de IA [17].

4.2. REQUISITO 2: SUPERVISIÓN HUMANA

La tecnología debe concebirse como una herramienta de asistencia, nunca de sustitución. Siguiendo las directrices de la OCDE sobre IA en la justicia y el artículo 14 del *AI Act*, la decisión final debe recaer siempre en un operador humano [16]. La IA debe actuar como un «borrador avanzado», automatizando la tarea de reescritura, pero dejando el control de calidad en manos del humano. Para que esta supervisión sea real, la interfaz del sistema debe resaltar los cambios realizados, permitiendo al supervisor verificar que la simplificación léxica no ha alterado la calificación jurídica de los hechos.

4.3. REQUISITO 3: SEGREGACIÓN FUNCIONAL

La metodología humana de la LF se basa en la separación de roles: un profesional adapta el texto y otro verifica su accesibilidad [2]. La IA debe emular esta arquitectura de control. No es suficiente con pedirle al modelo que «simplifique»; es necesario disociar el proceso en dos etapas: una instancia encargada de la redacción en lenguaje claro y una segunda instancia, independiente y con instrucciones adversarias, encargada de la validación normativa [26]. Esta segregación, propia de los sistemas de alta fiabilidad, reduce la tasa de error lógico, asegurando que la búsqueda de la sencillez no comprometa el rigor legal.

5. CONCLUSIÓN

La irrupción de la IA en el sector público nos sitúa ante un dilema: adoptar herramientas opacas que sacrifican garantías por eficiencia o construir sistemas que fortalezcan el acceso a la justicia. El análisis realizado evidencia que la adaptación a lectura fácil, exigida por la Ley 6/2022 [12], es inabordable manualmente, pero peligrosa si se delega en modelos de «Caja Negra». Para que la tecnología cumpla su promesa democratizadora, es necesario transitar hacia arquitecturas supervisadas que integren el anclaje factual y la validación humana como requisitos innegociables. Solo así garantizaremos que el derecho a entender de la ciudadanía se materialice sin socavar la seguridad jurídica que sustenta nuestra democracia.

6. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

7. REFERENCIAS

1. Alcaraz Varó, E., Hughes, B., & Gómez Martínez, M. A. (2014). *El español jurídico*. Ariel.
2. Asociación Española de Normalización. (2018). *Lectura fácil: Pautas y recomendaciones para la elaboración de documentos* (Norma UNE 153101:2018 EX). AENOR.
3. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM FAccT Conference on Fairness, Accountability, and Transparency* (pp. 610-623). ACM. <https://doi.org/10.1145/3442188.3445922>
4. Burgos Martínez, M. (2025). La paradoja de la aplicación del derecho a comprender en España: Una revisión normativa. *Sphera Publica*, 1(25), 112-141.
5. Centro de Referencia Estatal de Autonomía Personal y Ayudas Técnicas. (2023). *Lectura fácil (Ficha 3)* (Serie 10 fichas sobre la accesibilidad cognitiva por derecho). CEAPAT.
6. Consejo General del Poder Judicial. (2024). *Memoria anual 2024 sobre el estado, funcionamiento y actividades de los Juzgados y Tribunales*. CGPJ.
7. Cotino Hueso, L. (2023). Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida. *Revista Española de la Transparencia*, 16, 17-63. <https://doi.org/10.51915/ret.272>
8. Dahl, M., Maughan, J., Glazer, L., & Nyarko, J. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis*, 16(1), 64-93. <https://doi.org/10.1093/jla/lae003>
9. Hagan, M. (2018). A human-centered design approach to access to justice: Generating new prototypes and hypotheses for intervention to make courts user-friendly. *Indiana Journal of Law and Social Equality*, 6(2), 199-239.

10. Hagan, M. (2019). Participatory design for innovation in access to justice. *Daedalus*, 148(1), 120-127. https://doi.org/10.1162/daed_a_00544
11. High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. European Commission.
12. Jefatura del Estado. (2022, 1 de abril). Ley 6/2022, de 31 de marzo, de modificación del Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social, aprobado por el Real Decreto Legislativo 1/2013, de 29 de noviembre, para establecer y regular la accesibilidad cognitiva y sus condiciones de exigencia y aplicación. *Boletín Oficial del Estado*, (78), 43391-43401.
13. Laskowski, C., Rhee, S., & Shadiac, J. (2024). *Generative artificial intelligence and access to justice: Possibilities, concerns, best practices*. National Center for State Courts.
14. Lewis, P., Perez, E., Piktus, A., Petroni, F., Lewis, V., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474.
15. Montolío Durán, E. (Ed.). (2012). *Hacia la modernización del discurso jurídico*. Universitat de Barcelona.
16. OECD. (2019). *Equal access to justice for inclusive growth*. OECD Publishing. <https://doi.org/10.1787/597f5b7f-en>
17. Parlamento Europeo, & Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial). *Diario Oficial de la Unión Europea*, L, 2024/1689.
18. Plena Inclusión España. (2020). *10 consejos para mejorar el acceso a la justicia: Adaptación a lectura fácil de las directrices internacionales de la ONU*.
19. Plena Inclusión España. (2018). *Norma UNE 153101:2018 EX de lectura fácil: Pautas y recomendaciones para la elaboración de documentos*. Recuperado el 17 de diciembre de 2025, de <https://www.plenainclusion.org/publicaciones/buscador/norma-une-1531012018-ex-de-lectura-facil-pautas-y-recomendaciones-para-la-elaboracion-de-documentos/>
20. PwC Tax & Legal. (2022). *El derecho a entender: La comunicación clara como palanca de transformación jurídica*.
21. Real Patronato sobre Discapacidad. (2023). *España País Accesible: II Plan Nacional de Accesibilidad Universal 2023-2032*. Ministerio de Derechos Sociales y Agenda 2030.
22. Saggion, H. (2017). *Automatic text simplification*. Morgan & Claypool Publishers. <https://doi.org/10.2200/S00700ED1V01Y201602HLT032>
23. Simshaw, D. (2025). Interoperable legal AI for access to justice. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5425095>
24. Tu Web Accesible. (2022). *Ley 6/2022: La nueva accesibilidad cognitiva en España*. Recuperado el 17 de diciembre de 2025, de <https://tuwebaccesible.es/ley-6-2022-accesibilidad-cognitiva/>
25. Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasoor, S., Collins, Z., Burns, L., Hendricks, L. A., Wheeler, T., Isaac, W., & Gabriel, I. (2022). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214-229). ACM. <https://doi.org/10.1145/3531146.3533088>
26. Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Chang, B., Wang, Q., Jin, J., Fu, H., & Zhang, Y. (2023). *The rise and potential of large language model based agents: A survey* (arXiv:2309.07864). arXiv. <https://doi.org/10.48550/arXiv.2309.07864>

NUVORA: A TRUSTWORTHY AI FRAMEWORK FOR INTEGRATED DIGITAL HEALTH AND PATIENT WELL-BEING

NUVORA: UN MARCO DE IA FIABLE PARA LA SALUD DIGITAL INTEGRADA Y EL BIENESTAR DE LOS PACIENTES

Jaime SUÁREZ PÉREZ

Faculty of Economics and Business, University of Salamanca. Salamanca, 37007, Spain

✉ Jaimesuarez@usal.es

🌐 <https://orcid.org/0009-0007-4284-2926>

ABSTRACT: The rising prevalence of chronic diseases and polypharmacy presents clinical, economic, and ethical challenges to healthcare systems. Patients face fragmented pharmacological treatment management, financial constraints, and limited control over their health data. This paper presents Nuvora, a patient-centric digital health framework that integrates pharmacological and financial data through trustworthy artificial intelligence (AI) and blockchain-based governance. As a conceptual and architectural proposal derived from an innovation pitch, it emphasizes design principles, system architecture, and ethical foundations with-out empirical validation. The framework aims to support informed decision-making, improve treatment adherence, and enhance patient autonomy, aligning with European regulations on trustworthy AI and data protection.

Keywords: digital health; trustworthy ai; polypharmacy; patient-centric systems; blockchain; health data governance.

RESUMEN: El aumento de la prevalencia de las enfermedades crónicas y la polifarmacia plantea retos clínicos, económicos y éticos a los sistemas sanitarios. Los pacientes se enfrentan a una gestión fragmentada del tratamiento farmacológico, a restricciones económicas y a un control limitado sobre sus datos sanitarios. Este artículo presenta Nuvora, un marco de salud digital centrado en el paciente que integra datos farmacológicos y financieros mediante inteligencia artificial (IA) fiable y una gobernanza basada en *blockchain*. Como propuesta conceptual y arquitectónica derivada de una presentación de innovación, hace hincapié en los principios de diseño, la arquitectura del sistema y los fundamentos éticos sin validación empírica. El marco tiene como objetivo apoyar la toma de decisiones informadas, mejorar el cumplimiento del tratamiento y potenciar la autonomía del paciente, en consonancia con la normativa europea sobre IA fiable y protección de datos.

Palabras clave: salud digital; IA fiable; polifarmacia; sistemas centrados en el paciente; cadena de bloques; gobernanza de los datos sanitarios.

1. INTRODUCTION

Chronic diseases and polypharmacy represent a growing burden for healthcare systems worldwide [1-3]. Patients managing multiple long-term treatments often encounter fragmented care, limited coordination between healthcare providers, and insufficient tools to understand both the clinical and financial implications of their therapies. These challenges increase the risk of medication non-adherence, adverse drug interactions, and financial stress, ultimately affecting patient outcomes and quality of life.

Digital health solutions have emerged to address parts of these issues, such as medication reminders, telemedicine platforms, or financial health applications [4]. However, most existing tools remain siloed, focusing either on clinical data or financial aspects, without providing an integrated, patient-controlled view. Moreover, the increasing use of artificial intelligence in health raises concerns related to transparency, bias, data privacy, and regulatory compliance.

This paper introduces Nuvora, a digital health framework designed to integrate pharmacological and financial data within a single patient-centric ecosystem. Nuvora is conceived as a trustworthy AI-enabled platform that empowers patients to better understand, manage, and optimize their treatment plans. The contribution of this work lies not in empirical validation, but in the structured presentation of the concept, architecture, and ethical foundations of the proposed system, aligned with current European regulatory and societal expectations.

2. NUVORA: CONCEPT AND ARCHITECTURE

2.1. CONCEPTUAL OVERVIEW

Nuvora is designed as a holistic digital companion for patients with chronic conditions and complex medication regimens. Its core objective is to support informed decision-making by combining health-related data with financial insights, placing the patient at the center of the system. Unlike traditional health applications, Nuvora does not aim to replace healthcare professionals but to complement clinical decision-making through personalized, transparent, and explainable recommendations.

The framework assumes that patient well-being is influenced not only by clinical factors but also by economic constraints and behavioral patterns. By addressing these dimensions simultaneously, Nuvora seeks to reduce treatment-related risks and enhance patient engagement.

2.2. DATA INTEGRATION LAYER

The system integrates heterogeneous data sources, including:

- Pharmacological data (prescriptions, dosage, treatment duration).
- Patient-reported information (adherence, side effects).
- Financial data related to medication costs and healthcare expenses.

Data integration is designed following privacy-by-design principles, ensuring that only the minimum necessary data are processed and that patients retain control over data access and sharing permissions.

2.3. AI-DRIVEN DECISION SUPPORT

Artificial intelligence within Nuvora is used to provide decision support rather than automated clinical decisions. The AI layer focuses on:

- Identifying potential risks related to polypharmacy.
- Detecting adherence patterns.
- Offering personalized alerts and recommendations.

The system emphasizes explainability, allowing users to understand the rationale behind recommendations. This approach aligns with the concept of trustworthy AI, prioritizing transparency, human oversight, and accountability over purely predictive performance.

2.4. TRUST AND GOVERNANCE THROUGH BLOCKCHAIN

To reinforce trust and data governance, Nuvora incorporates blockchain-based mechanisms for consent management and data traceability. Blockchain is not used for data storage but as a governance layer that records access permissions and consent transactions. This design enables auditable data flows and strengthens patient autonomy over personal health information.

3. TRUSTWORTHY AI AND ETHICAL DESIGN

Trustworthy AI is a foundational principle of the Nuvora framework. The system is designed in alignment with European ethical guidelines and emerging regulatory frameworks, particularly those addressing high-risk AI systems in healthcare. Key ethical considerations include:

- *Transparency*: AI-driven recommendations are explainable and accessible to users.
- *Privacy and Data Protection*: Compliance with data protection regulations is ensured through anonymization, access control, and decentralized governance.
- *Human Oversight*: Healthcare professionals remain central to clinical decision-making.
- *User Empowerment*: Patients can actively manage consent and data sharing preferences.

By embedding these principles into the system architecture, Nuvora aims to mitigate common risks associated with AI in healthcare, such as loss of trust, opaque decision-making, and data misuse.

4. DISCUSSION AND FUTURE WORK

As a conceptual framework derived from a pitch, Nuvora presents several strengths, including its integrated perspective on health and financial well-being and its emphasis on trustworthy AI. However, the approach also faces limitations. The absence of large-scale empirical validation and real-world pilots currently restricts claims regarding clinical effectiveness or economic impact.

Future work will focus on pilot studies in collaboration with healthcare institutions, technical validation of the AI components, and usability assessments with end users. These steps are essential to evaluate the practical feasibility and real-world benefits of the proposed framework.

5. CONCLUSIONS

This short paper has presented Nuvora as a patient-centric digital health framework that integrates pharmacological and financial data using trustworthy AI and blockchain-based governance. By translating an innovation pitch into an academic format, the paper highlights the conceptual and ethical foundations of the system rather than empirical outcomes. Nuvora contributes to ongoing discussions on integrated digital health solutions and underscores the importance of trust, transparency, and patient empowerment in the design of future healthcare technologies.

6. AGRADECIMIENTOS

Esta investigación forma parte del Proyecto de Cátedra Internacional sobre Inteligencia Artificial Fiable y Reto Demográfico dentro de la Estrategia Nacional de Inteligencia Artificial (ENIA), en el marco del Plan Europeo de Recuperación, Transformación y Resiliencia. Referencia: TSI-100933-2023-0001. Este proyecto está financiado por la Secretaría de Estado para la Digitalización y la Inteligencia Artificial y por la Unión Europea (Next Generation).

7. REFERENCES

1. PRIMA-eDS Consortium. (2017). *Final report summary - PRIMA-eDS (Polypharmacy in chronic diseases: Reduction of Inappropriate Medication and Adverse drug events in elderly populations by electronic Decision Support)*. European Commission CORDIS. <https://cordis.europa.eu/project/id/305388/reporting>
2. Ministerio de Sanidad. (2025). *Casi el 55% de la población española mayor de 15 años tiene una enfermedad crónica*. Gobierno de España, Gabinete de Prensa. <https://www.sanidad.gob.es/en/gabinete/notasPrensa.do?id=6771>
3. García-Molina Sáez, C., Romeo-Sánchez, Z., Montero-Salinas, A., Martínez-Martínez, F., & Gastelurrutia, M. A. (2012). Prevalencia de polimedicación y riesgo vascular en la población mayor de 65 años. *Atención Primaria*, 44(4), 202-209. <https://doi.org/10.1016/j.aprim.2011.05.004>
4. Bennie, M., Santa-Ana-Tellez, Y., Galistiani, G. F., Mair, A., Mueller, T., Elseviers, M., Wettermark, B., & Henman, M. (2024). The prevalence of polypharmacy in older Europeans: A multinational database study of general practitioner prescribing. *British Journal of Clinical Pharmacology*, 90(10), 2124-2136. <https://doi.org/10.1111/bcp.16113>

AQUILAFUENTE, 394

Esta obra colectiva constituye un hito en la literatura científica actual al abordar, desde una perspectiva técnica y humanista, uno de los desafíos más críticos de la Europa del siglo XXI: la supervivencia del mundo rural a través de la digitalización. El libro compendia las investigaciones presentadas en el *I Simposio de Inteligencia Artificial aplicada al reto demográfico*, consolidándose como un manual de referencia para investigadores, tecnólogos y gestores públicos.

El volumen se estructura en torno a la idea de la «IA Fiable», articulando soluciones en cuatro áreas fundamentales: 1) Sostenibilidad y sector primario: se presentan estudios pioneros sobre viticultura de precisión ante el cambio climático y herramientas como *DeepForest* para la gestión de biomasa. Se propone una modernización del campo que no solo es productiva, sino ecológicamente responsable. 2) Salud y bienestar rural: se aborda la teleasistencia avanzada y uso de modelos de lenguaje para reducir la brecha de comunicación médica. Destaca la visión del «envejecimiento en el hogar» (*ageing in place*) apoyado por sensores inteligentes que respetan la privacidad de las personas. 3) Economía y territorios inteligentes: proyectos como *Ruralink* (empleo juvenil) y *DataMarket* (comercio local), proponen un modelo de desarrollo donde la IA democratiza el acceso a la analítica de datos, tradicionalmente reservada a grandes urbes. 4) Marco ético y legal: se analiza la gobernanza de estas tecnologías bajo el nuevo Reglamento de IA de la UE, enfatizando la importancia de la «lectura fácil» y la transparencia algorítmica.

Este libro es un manifiesto tecnológico pues defiende que la IA no debe ser un factor de exclusión, sino un puente de cohesión territorial. Es una lectura esencial para entender cómo la innovación puede devolver la vitalidad a las zonas despobladas, garantizando que el progreso técnico sea, ante todo, un progreso social.



Colabora

