

DEBATE/DEBATE: MÁS ALLÁ DEL BIG DATA: IA GENERATIVA Y LLMS COMO NUEVAS TECNOLOGÍAS DIGITALES PARA EL ANÁLISIS DE LA REALIDAD SOCIAL / BEYOND BIG DATA: GENERATIVE AI AND LLMS AS NEW DIGITAL TECHNOLOGIES FOR THE ANALYSIS OF SOCIAL REALITY

Editorial: Inteligencia artificial generativa, grandes modelos de lenguaje (LLMs) y analítica aumentada vs. *big data* y ciencia de datos: Nuevas avenidas para la investigación social

Editorial: Generative Artificial Intelligence, Large Language Models (LLMs), and Augmented Analytics vs. Big Data and Data Science

Estrella Gualda

Universidad de Huelva (UHU, ESEIS/COIDESO/CISCOA-Lab), España
estrella@uhu.es

Recibido/Received: 9-6-2025

Aceptado/Accepted: 14-10-2025



RESUMEN

Este artículo examina comparativamente la evolución de campos y tecnologías como la inteligencia artificial, la inteligencia de negocios, la ciencia de datos, los big data, los grandes modelos de lenguaje (LLMs), la inteligencia generativa y la analítica aumentada. A partir de los documentos «más citados» y «candidentes» de WoS, analiza redes de co-ocurrencias de términos, visualizando un mapa de conocimiento donde se observan temas o conceptos clave de esta bibliografía y sus conexiones. En el último lustro se aprecia una tendencia a la reducción relativa de las publicaciones científicas sobre la inteligencia de negocios, *big data* o ciencia de datos en comparación con las publicaciones sobre LLMs —como ChatGPT—, inteligencia generativa y analítica aumentada. Se anticipa una interesante diversidad de oportunidades e impactos para las ciencias sociales, como muestran con detalle los artículos incluidos en esta sección de Debate de la Revista *CENTRA de Ciencias Sociales*.

PALABRAS CLAVE: Inteligencia artificial; inteligencia de negocios; ciencia de datos; macrodatos; modelos grandes de lenguaje; inteligencia generativa; analítica aumentada; Web of Science; mapas de conocimiento; publicaciones científicas; ciencias sociales.

CÓMO CITAR: Gualda, E. (2026). Editorial: Inteligencia artificial generativa, grandes modelos de lenguaje (LLMs) y analítica aumentada vs. *big data* y ciencia de datos: Nuevas avenidas para la investigación social. *Revista Centra de Ciencias Sociales*, 5(1), 157-172. <https://doi.org/10.54790/rccs.150>

English version can be read on <https://doi.org/10.54790/rccs.150>

ABSTRACT

This article comparatively examines the evolution of fields and technologies such as artificial intelligence (AI), business intelligence (BI), data science (DS), big data (BD), large language models (LLMs), generative intelligence, and augmented analytics. Based on the “most cited” and “hot papers” in Web of Science (WoS), it analyzes co-occurrence networks of cited terms, visualizing a knowledge map that highlights key concepts and their connections. Over the past five years, there has been a relative decline in scientific publications on BI, BD, and DS, contrasted with the growing focus on LLMs —such as ChatGPT— generative artificial intelligence, and augmented analytics. This shift marks a significant transformation and opens up a range of opportunities and impacts for the social sciences, as detailed in the articles included in the Debate section of Revista CENTRA de Ciencias Sociales.

KEYWORDS: Artificial intelligence; business intelligence; data science; big data; large language models; generative intelligence; augmented analytics; Web of Science; knowledge maps; scientific publications; social sciences.

1. De los *big data* y la ciencia de datos a la inteligencia artificial generativa, los grandes modelos de lenguaje y la analítica aumentada

Aunque estamos quizás iniciando una senda emergente que no sabemos aún hacia dónde va a llevar a las ciencias sociales, se anticipan grandes transformaciones en algunos parámetros clave respecto a nuestros procesos de trabajo intelectual y a nuestra forma de enfocar la investigación a consecuencia de los grandes avances que llegan de la mano de tecnologías y herramientas como las de los grandes modelos de lenguaje (LLMs), la inteligencia artificial generativa y la analítica aumentada. Las recientes innovaciones diversifican y amplían los impactos para las ciencias sociales que la inteligencia artificial, la inteligencia de negocios, los *big data* y la ciencia de datos ya aportaban.

El impacto social o transformaciones derivadas de estas innovaciones, y especialmente de la IA generativa, ya comenzó (Saetra, 2023), y abarca a múltiples áreas y sectores, como, por ejemplo, los de la educación (Walter, 2024; Chiu, 2024) o la sanidad, donde se llega a hablar de un cambio de paradigma en aspectos relativos a las imágenes en medicina (Pinto-Coelho, 2023).

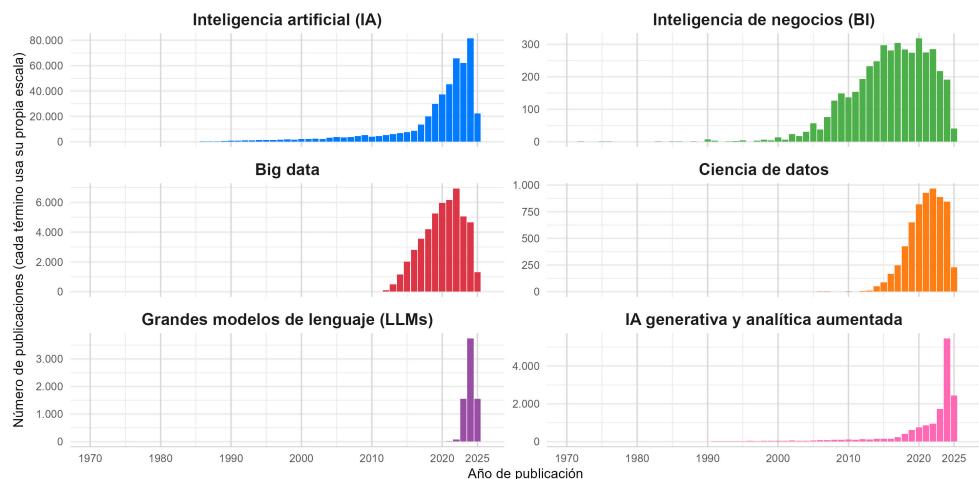
Intuimos, por nuestro conocimiento de las ciencias sociales, y como igualmente anticipan la lectura de los textos de Gandler (2026) y Gómez Espino (2026) incluidos en esta sección de DEBATE de la *Revista CENTRA de Ciencias Sociales*

[<https://www.centracs.es/revista>], que asistiremos en los próximos años a la eclosión de estrategias de investigación que se apoyan en alguno o varios de los adelantos ligados a las áreas citadas, en combinación con incursiones más o menos profundas en el *prompting* o la ingeniería de *prompts* (Walter, 2024), que ya está permitiendo a las ciencias sociales (y a otras como las de la salud, Meskó, 2023) beneficiarse de estos avances sin el hándicap, previamente existente, de tener que contar con un dominio avanzado de materias como la programación. Esta eclosión de investigaciones que saquen partido de estas innovaciones no es incompatible con aproximaciones de corte más clásico o mixtas.

La trayectoria seguida en la investigación científica en estos campos, tecnologías y herramientas se vislumbra con nitidez explorando comparativamente sus publicaciones, lo que hacemos a continuación de forma sintética por la limitación de páginas que tenemos. Con el objetivo de hacer una comparación introductoria, se han elaborado una serie de gráficos de barras con apoyo de R, que muestran la evolución de las publicaciones que contienen, sea en sus títulos o en sus resúmenes, los términos extraídos en seis búsquedas bibliográficas diferentes: (1) «artificial intelligence OR AI», (2) «business intelligence», (3) «big data», (4) «data science» y (5) «large language models OR LLMs». Por último, para identificar aspectos más emergentes, que conectan con los avances en la inteligencia artificial analítica y la analítica aumentada, se llevó a cabo una sexta búsqueda que integró términos clave como (6) «generative AI», «generative artificial intelligence», «diffusion models», «GANs», «augmented analytics», «automated insights», «natural language generation» o «AI-driven analytics», términos genéricos y específicos que permiten hacernos una idea de las investigaciones en marcha en esta área para, posteriormente, valorar algunos de los aspectos que pueden contribuir a enriquecer el trabajo de las ciencias sociales. Las búsquedas se hicieron en la base de datos bibliográfica Web of Science (WoS, <https://www.webofscience.com/wos/>), con un margen temporal que comenzaba desde el inicio de la serie de publicaciones hasta finales de mayo de 2025. Se filtraron las publicaciones según el dominio de «Cien-cias Sociales» («Research Domain» en WoS).

La figura 1 presenta recuentos anuales absolutos de publicaciones para cada búsqueda. Para facilitar la lectura de las trayectorias temporales, cada panel utiliza un eje vertical independiente (escala ajustada al rango de su propia serie). Se advierte que la altura de las barras no es comparable entre paneles; solo lo es dentro de cada panel. Dadas las diferencias de volumen entre áreas, optamos por este diseño de paneles con ejes independientes para percibir con claridad la evolución en cada caso, sin recurrir a escalas logarítmicas. No se aplicó ninguna normalización a las series: las barras muestran recuentos absolutos por año en cada panel.

Figura 1
Evolución de las publicaciones en Ciencias Sociales (por campos y tecnologías) - WoS



Fuente: elaboración propia realizada con R, a partir de datos de WoS.

Este gráfico tiene una versión interactiva. Tanto el gráfico interactivo como el código y datos empleados para su realización pueden consultarse en el siguiente enlace: <https://doi.org/10.5281/zenodo.17298490> (Gualda, 2025).

Esta estrategia comparativa nos permite hacernos una muy buena idea de la evolución seguida, así como de los principales hitos, para profundizar posteriormente en cuáles son los términos clave que identifican estos campos y tecnologías y cómo conectan entre sí. Nos permite también aproximarnos a algunas de sus ramificaciones más relevantes, a través de un análisis de redes o de co-ocurrencia de conceptos clave (mapas de conocimiento), ya empleado previamente en otros estudios bibliométricos (Cobo *et al.*, 2011).

Aunque se extrajeron todos los datos de WoS sobre publicaciones sin limitación temporal, después de un primer preanálisis, para la elaboración de los gráficos, hemos optado por visualizar la serie temporal comparativa desde 1970 hasta finales de mayo de 2025 (cuando se cierra la escritura de este texto). En el periodo previo a los setenta, sobre todo desde los cincuenta, el campo de la inteligencia artificial se encontraba ya más que presente, manejándose desde los inicios la idea de que era posible crear máquinas capaces de simular el pensamiento humano o de hacer cosas que «se considerarían inteligentes si las hiciera un ser humano» (Skinner, 2012, p. 3). Incluso en las ciencias sociales se rastrean ya algunas exploraciones sobre el uso de las computadoras para el análisis estadístico y la simulación de fenómenos sociales desde hace tiempo, con la aparición de paquetes emblemáticos como Spss en 1968. El texto de Gandler (2026) en esta sección de DEBATE aborda de manera detallada esta evolución histórica para las ciencias sociales y las humanas en sus conexiones con las tecnologías digitales.

Por otra parte, en torno a los sesenta, ya se pueden documentar los primeros pasos de la inteligencia de negocios (Luhn, 1958), que incorpora algunos de sus primeros conceptos y sistemas para organizar y analizar datos empresariales con el fin de mejorar la toma de decisiones. Igualmente ocurre con respecto a la creación de bases de datos relacionales que van sentando las bases para organizar datos (Codd, 1970), como precedentes de sistemas modernos.

Comparando la evolución de las publicaciones en los seis campos y tecnologías citadas, se aprecia que las publicaciones sobre inteligencia artificial que conectan con las ciencias sociales comienzan mucho antes que otras, mostrando un crecimiento muy significativo a partir de 2010 aproximadamente y muy intenso en el último lustro (figura 1). El interés por la inteligencia artificial en publicaciones ligadas a las ciencias sociales es sostenido en el tiempo, aunque en sus primeras décadas su crecimiento fue lento. Por otra parte, si bien se identifican publicaciones sobre la inteligencia de negocios un poco más tarde que sobre inteligencia artificial, ya destacan gradualmente desde inicio de los noventa y sobre todo desde 2005 se encuentra un crecimiento sostenido que es muy apreciable también respecto al volumen de publicaciones. Recordamos que la evolución que estamos trazando se refiere a la evolución de bibliografía sobre estos campos y tecnologías filtrando las publicaciones que en WoS aparecen clasificadas en el dominio de «Ciencias Sociales».

Frente al inicio más temprano de las áreas de inteligencia artificial e inteligencia de negocios, las áreas de *big data* y ciencia de datos, que frecuentemente se dan la mano, comienzan a tener más interés para las ciencias sociales ya a partir de la segunda década de este siglo con una eclosión de trabajos primero más importantes respecto a los *big data* desde 2010–2012, época en la que también se aprecia un despuente paralelo, pero algo más tímido comparativamente de las publicaciones encuadradas en la ciencia de datos, muy conectadas. Lo que parece apreciarse en el último lustro es una caída relativa de documentos que aluden específicamente a los *big data*, frente al incremento de las publicaciones sobre ciencia de datos.

Aspectos clave de esta evolución, ya en el siglo XXI, son los avances y el desarrollo del aprendizaje profundo (DL) desde 2010, revolucionando la IA, y de los grandes modelos de lenguaje (LLMs), un poco más adelante, con la introducción de tecnologías como la del ChatGPT o similares (Meskó, 2023; Saetra, 2023; Cooper, 2023). Avances que vienen de la mano del aprendizaje profundo en aspectos como el reconocimiento de imágenes, voz y procesamiento de lenguaje natural que tienen profundos impactos en las ciencias sociales. Paralelamente, la creación de modelos de lenguaje con fundamento en arquitecturas como los *transformers*, que pueden generar y comprender texto muy rápidamente, supone un nuevo impacto para las ciencias sociales. Sobre este particular, centrado en los procesos de codificación cualitativa temática, se enfoca el artículo de Gómez Espino (2026) en esta sección de DEBATE, que presta atención a los modelos grandes de lenguaje y sus aportaciones cualitativas al crear la herramienta *Social Verbatim*, subrayando la importancia de la colaboración humano-máquina.

Aunque las publicaciones sobre grandes modelos de lenguaje o inteligencia generativa y analítica aumentada son campos de estudio y tecnologías que se han desarrollado muy recientemente, el crecimiento exponencial de publicaciones que apreciamos en los gráficos de la figura 1 parece sugerir que van a ser áreas en desarrollo intenso en las ciencias sociales en los próximos tiempos. Los LLM o modelos grandes de lenguaje y el conocido como «prompting» pueden suponer un impulso adicional para el desarrollo de una diversidad de tareas asociadas a la investigación en general o la que es más común en ciencias sociales, estimulando la imaginación sociológica (Mills, 1959). La potencialidad de contar con la inteligencia artificial como «asistente técnico» en sus versiones más recientes (LLMs, generativa y analítica) puede ayudar a las ciencias sociales a sortear algunos hándicaps técnicos precedentes y enriquecer el trabajo en un abanico de aspectos como la búsqueda de la bibliografía, la implementación de tareas como la codificación y clasificación de documentos, la minería de textos, la realización de análisis sofisticados con gran rapidez a partir de diversidad de tipos de datos (texto, imagen, sonido, vídeo), la mejora en las visualizaciones, y un largo etcétera.

2. Mapas de conocimiento y clústeres significativos referidos a literatura sobre inteligencia artificial clásica y moderna, inteligencia de negocios, *big data* y ciencia de datos

Una vez conocida *grosso modo* la evolución de las publicaciones en las ciencias sociales según campos y tecnologías, nos vamos a detener ahora en observar las redes de coocurrencias y clústeres entre términos, pudiendo esbozar «mapas de conocimiento» (Cobo *et al.*, 2011) sobre el campo de la inteligencia artificial (clásica y moderna) y líneas de trabajo como las de la inteligencia de negocios, *big data* y ciencia de datos. Por la limitación de páginas, se llevará a cabo un análisis global muy sintético, si bien se pueden apreciar algunos detalles más concretos en las figuras que hemos incluido. Prestaremos atención para esta identificación solo a los documentos «más citados» y «candentes» de WoS, visualizando tanto temas o conceptos clave de esta bibliografía como sus conexiones.

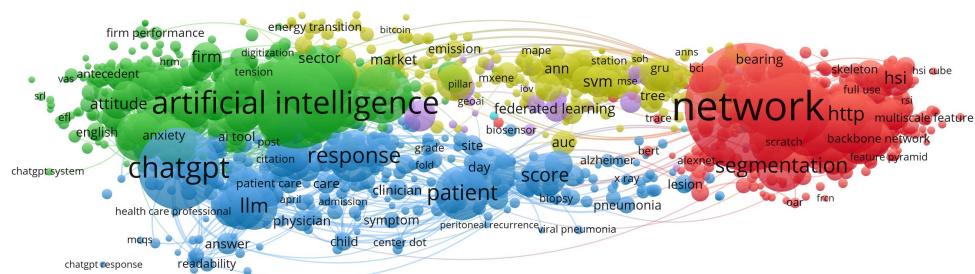
A estos efectos, del conjunto de artículos sobre los que hemos trazado la evolución, hemos extraído los clasificados en WoS como los más citados y los candentes («hot») en las seis búsquedas realizadas (inteligencia artificial, inteligencia de negocios, *big data*, ciencia de datos, grandes modelos de lenguaje e IA generativa y analítica aumentada). A partir de aquí, hemos agregado estos datos fusionando en R las seis bases de datos obtenidas en WoS, se han eliminado los documentos duplicados y se adaptó el archivo resultante para la importación de los datos en el programa de análisis bibliométrico VOSviewer (Van Eck y Waltman, 2010), a partir del que se han calculado métricas de relevancia, para posteriormente visualizar o mapear los clústeres más significativos.

La figura 2 muestra un mapa de conocimiento que contiene los términos más relevantes en las obras más citadas en la bibliografía. El análisis de redes con VOSviewer

ha permitido identificar varios clústeres que agrupan términos conectados entre sí y que permiten representar aspectos centrales en esa área temática de la literatura. Los términos se extrajeron del título y resumen de cada documento. La visualización muestra solo términos muy relevantes, lo cual nos da una pista clara en cada subgrafo identificado por colores de las principales áreas de investigación. Las interconexiones permiten observar los aspectos más conectados en la literatura. En el análisis sintético que sigue nos referimos a algunos de los términos con más influencia a tenor del indicador «fuerza de enlace total» (*Total link strength*), métrica que revela los pilares de cada clúster. El tamaño de cada nodo en el clúster se asocia igualmente a la centralidad de cada término. La proximidad entre los términos sugiere igualmente una mayor relación o coocurrencia.

Figura 2

Mapa de conocimiento sobre campos y tecnologías asociados a la inteligencia artificial



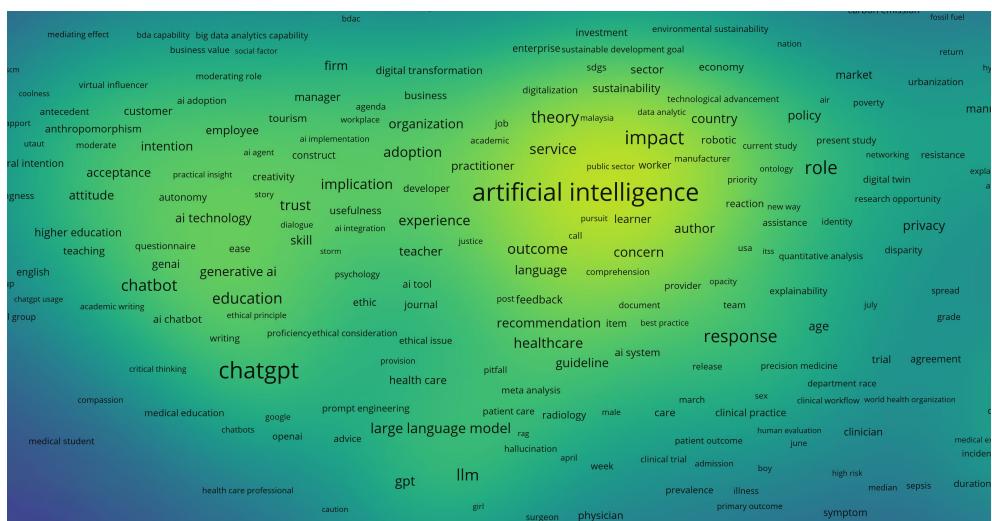
Fuente: elaboración propia a partir de VOSviewer. Datos extraídos de WoS y procesados en R.

2.1. Inteligencia artificial, sociedad e interacción humano-máquina (clúster verde)

El clúster verde, centrado en la inteligencia artificial, es uno de los conglomerados clave en este mapa de conocimiento, revelando la importancia de la literatura sobre la adopción, percepción, uso y el impacto de la IA, así como de sus tecnologías relacionadas en diversos contextos sociales y organizacionales. O con aspectos como el impacto en la experiencia del cliente, la gestión empresarial y la aceptación del usuario. No es extraño que algunos de los términos clave más relevantes¹ sean los de «adoption», «acceptance», «AI adoption», «AI technology», «AI use», «usage intention» o «technology acceptance model (TAM)», que conectan con investigaciones sobre los factores psicológicos y de comportamiento que influyen en este proceso —por ejemplo, la utilidad («usefulness») y la facilidad de uso de las tecnologías («ease»)—. Algunos de estos términos y otros muchos se aprecian en el mapa de color representando en la figura 3.

Figura 3

Mapa de calor o densidad del clúster verde



Fuente: elaboración propia a partir de VOSviewer. El color amarillo representa una mayor densidad o concentración de publicaciones relacionadas con el tema marcado.

Se recoge también en el clúster la bibliografía que estudia la interacción humano-IA/robot, lo que se aprecia en términos como «human AI interaction», «human robot interaction (HRI)», «human machine interaction» o «robot». De forma complementaria, se encuentra la investigación sobre emociones y actitudes humanas («anxiety», «AI anxiety», «fear», «negative emotion», «positive emotion»), y sobre las características de los robots que influyen en estas interacciones, incluyendo el antropomorfismo («anthropomorphism») y la presencia social («social presence»), aspectos muy presentes en la bibliografía.

Este amplio clúster también pone énfasis en el impacto de la IA en la educación y el aprendizaje («education», «higher education», «student», «teacher», «academic writing» y «language learning»). La literatura explora cómo la IA afecta los procesos de enseñanza-aprendizaje y los desafíos éticos asociados, como el «plagiarism», la privacidad, consideraciones éticas y otras. También se documentan impactos en el ámbito de los negocios, la gestión y la transformación digital, donde la IA redefine el sector empresarial, en un sentido amplio. Lo atestiguan la relevancia de términos como «business», «digital transformation», «innovation», «sustainability», «job», «employment», «supply chain» o «customer», propios de una diversidad de sectores como la hostelería, el turismo o los servicios. Este clúster, aunque no es el único, se distingue por su énfasis en los LLM y la IA generativa, lo que se aprecia en el uso frecuente en la bibliografía de términos como «chatbot» y «ChatGPT», reflejando la experimentación en estos campos, lo que contrasta con publicaciones más tradicio-

nales basadas en modelos predictivos. Se enfatiza igualmente en este conglomerado la investigación y metodología rigurosas, siendo comunes términos en esta área («empirical study», «qualitative study», «SEM», etc.), lo que subraya la importancia de la validación empírica y la construcción de teorías necesarias para comprender y gestionar eficazmente los impactos de la IA.

2.2. Redes neuronales profundas: Visión por computadora, detección y segmentación de imágenes (clúster rojo)

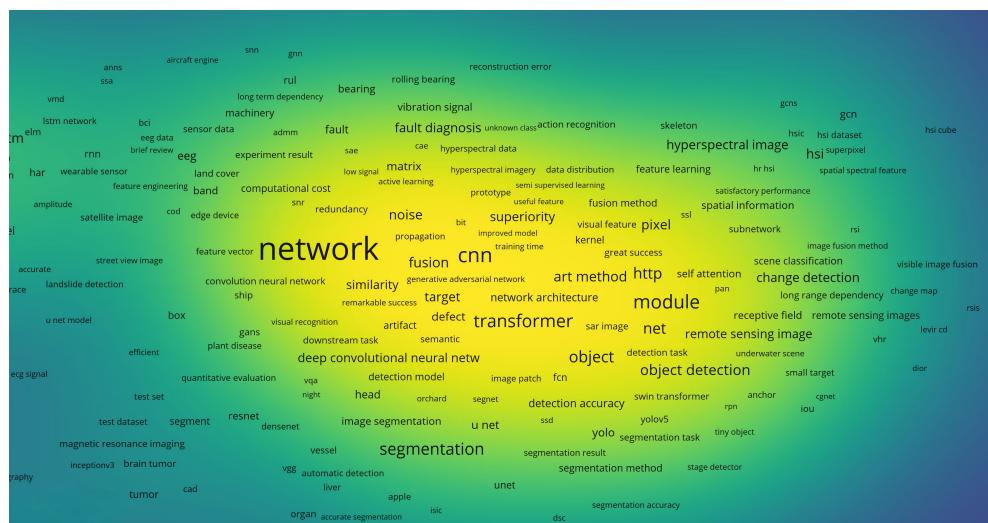
El clúster engloba conceptos centrales de arquitecturas como las redes neuronales convolucionales (CNNs) y los transformers. Su bibliografía subraya aplicaciones clave en la investigación reciente de inteligencia artificial, de gran interés y potencial en las ciencias sociales, enfocándose en la visión por computadora, detección de objetos y segmentación de imágenes.

Las publicaciones de este clúster representan un epicentro en la investigación actual en inteligencia artificial, de gran interés en nuestro campo por el potencial, entre otros, del procesamiento y análisis de imágenes. El término «network» es el de mayor conectividad del clúster. Otros términos clave propios de estos campos y sobre los que gira la bibliografía (figura 4) son «module», «segmentation», «detection», «fault diagnosis», «target detection», «neural», «image fusion», «layer», «experimental result», «code», «github.com», «superiority». Son igualmente centrales términos relativos a algoritmos como «CNN», «LSTM» y «SVM».

A tenor de la bibliografía, la investigación se enfoca en el diseño de arquitecturas eficientes, la mejora continua del rendimiento y la aplicabilidad de los modelos en diversos escenarios del mundo real. Esto incluye la extracción y fusión de características, y la búsqueda de modelos «efficient» con «excellent performance». Los documentos que caracterizan este clúster parecen destacar por su fuerte compromiso con la experimentación, la validación empírica y la colaboración a través de la compartición de código, buscando resolver desafíos como el desequilibrio de clases y el sobreajuste para mejorar la «generalization ability».

Figura 4

Mapa de calor o densidad del clúster rojo



Fuente: elaboración propia a partir de VOSviewer. El color amarillo representa una mayor densidad o concentración de publicaciones relacionadas con el tema marcado.

2.3. IA y LLM en el ámbito de la salud: del diagnóstico, a la toma de decisiones y sus aspectos éticos (clúster azul)

Este es un conglomerado muy grande y diverso, donde confluye la investigación en inteligencia artificial (IA) y el sector de la salud y la medicina. Se orienta en gran medida a aplicaciones de la IA y la ciencia de datos en este ámbito, destacando el uso de LLMs como el ChatGPT. Algunos términos clave son precisamente «ChatGPT», «LLM», «healthcare», «medicine», «patient», «clinical», «diagnosis», «disease», «symptom», «physician», «clinical decision», «symptom», «drug discovery», «pneumonia» y «COVID».

La investigación en esta área se centra en el diagnóstico y detección de enfermedades (como cáncer, neumonía y Alzheimer) utilizando IA para mejorar la precisión y eficiencia, apoyándose en técnicas de imagen médica como radiografías de tórax y tomografías computarizadas. También explora el apoyo a la toma de decisiones clínicas y el manejo de datos de pacientes. Además, el clúster aborda las consideraciones éticas, sociales y regulatorias de la IA en salud, discutiendo principios como la rendición de cuentas, la transparencia y la equidad, así como desafíos como el sesgo algorítmico y la desinformación. La metodología de investigación es fundamental aquí, subrayando la necesidad de una base de evidencia sólida a través de ensayos clínicos y revisiones sistemáticas para la adopción segura y justa de la IA en la práctica médica. En esencia, este clúster representa el ecosistema integral de la IA en el sector sanitario, cubriendo

tanto las capacidades tecnológicas (diagnóstico, apoyo clínico, procesamiento de datos) como las implicaciones éticas y sociales críticas para su implementación.

2.4. IA distribuida, *blockchain*, privacidad y optimización (clúster púrpura)

Este conglomerado refleja la intersección de la inteligencia artificial con tecnologías distribuidas, seguridad, privacidad y optimización. Los términos que lo identifican revelan que se enfoca en la implementación segura y eficiente de la IA en sistemas distribuidos o descentralizados, priorizando la seguridad, confidencialidad y eficiencia para aplicaciones a gran escala. Algunos términos clave son: «blockchain technology», «distributed ledger technology (DLT)», «federated learning», «edge computing», «NFT» o «cryptocurrency» en el ámbito de las tecnologías distribuidas. Otros términos aluden a la seguridad y privacidad («data privacy», «privacy policy», «privacy preserving», «privacy leakage», «privacy protection», «attack», «cybersecurity», «distributed attack» o «malicious attack»).

Respecto a la optimización y eficiencia, se identifican en el mapa algunos como «optimization», «algorithm», «resource allocation» o «energy efficiency». Y en el área de aplicaciones que emergen pueden citarse algunas como «metaverse», «digital twin», «smart city», «smart manufacturing», «internet of vehicles» o «cyber physical system». También se destacan otros como «content generation», «recommender system», «deep reinforcement learning», «graph neural network», «data management», «data sharing» o «simulation». La investigación aborda la operación eficiente y segura en contextos como la federación de datos, la gestión de la cadena de suministro, la infraestructura inteligente y el metaverso, enfatizando la protección de la información y la prevención de ataques con la «simulation» como metodología clave para validar soluciones.

2.5. *Machine learning*, analítica predictiva e inteligencia de negocios (clúster amarillo)

Este clúster mapea conceptos relativos a técnicas y modelos avanzados de *machine learning* y aprendizaje profundo aplicados al análisis de datos y la predicción. Incluye arquitecturas como «lstm», «svm» (Support Vector Machine), «random forest», «xgboost» y «ann» (artificial neural network). Algunos términos clave que aparecen en la bibliografía tienen que ver con métricas de validación como «auc» y «rmse». El clúster también abarca la «inteligencia de negocios» (BI) y la aplicación de estos modelos al sector empresarial o de mercado («market», «firms», etc.).

Se aprecia globalmente, como balance, que diferentes clústeres del mapa manifiestan una fuerte interconexión. Así, el clúster verde (inteligencia artificial) parece actuar como eje central, conectando directamente con el azul (IA y LLMs en salud), el amarillo (*machine learning* y BI), y el rojo, con investigaciones más novedosas sobre redes neuronales profundas y visión por computadora, evidenciando a la IA como un núcleo que une estas áreas aplicadas y técnicas.

3. Avenidas para la investigación social: a modo de conclusión

Si bien el breve y genérico recorrido que hemos hecho a través de la bibliografía contando con una única fuente (WoS) solo permite dar algunas pinceladas de las intensas transformaciones en curso, las nuevas tecnologías o herramientas ligadas a la inteligencia artificial que nos llegan a las ciencias sociales abren un abanico inmenso de oportunidades para la investigación, el análisis y la intervención en las diferentes disciplinas que las componen. Aunque trazar una evolución sintética no era una tarea sencilla, nos ha permitido introducir y enmarcar esta sección de DEBATE de la *Revista CENTRA de Estudios Sociales* y aproximarnos a algunos de los temas que más interés han suscitado en las ciencias sociales en diferentes momentos históricos, así como observar campos emergentes en el área de la inteligencia artificial a los que prestar especial atención. Un análisis más exhaustivo al respecto se encuentra en el texto de Gendler (2026) que se incluye en este número.

Nos referimos ahora especialmente, dentro del amplio campo de la IA, a los desarrollos relativos a los LLM y a la IA generativa y analítica, como modelos (frecuentemente multimodales, Mésko, 2023) que tienen como característica poder generar contenido nuevo y original, y no solo analizar o clasificar datos. La capacidad de generación o creación de texto, imágenes, audio, vídeo o código es un rasgo muy avanzado al que no se puede dejar de prestar atención.

Esta capacidad generativa, junto a sus ventajas, está generando frecuentemente grandes preguntas, debates y críticas que se refieren a aspectos negativos o riesgos como los posibles daños que pueden surgir cuando usamos la IA para generar contenidos (Saetra, 2023; Harrer, 2023), como es el caso de la misma multiplicación de la capacidad de generar desinformación, odio, «deepfakes» o contenidos dañinos o de dudosa veracidad con ayuda de una IA generativa que funcione sin supervisión humana (Harrer, 2023).

Otras dimensiones críticas tienen que ver con sesgos inherentes a los algoritmos (Varsha, 2023); críticas relativas al plagio de obras académicas, artísticas u otras, preocupaciones éticas o incluso el riesgo de infracción de los derechos de autor, que se añaden a otros problemas conectados entre sí y relacionados con la privacidad, anonimización y seguridad de los datos, la protección, la evitación del control de los datos por manos privadas, la ética o la gobernanza de la IA (Murdoch, 2021; Mügge, 2024). Otras críticas sustantivas, como se avanzó en relación con el clúster verde, aluden a la tendencia de antropomorfizar la inteligencia artificial (Ryan, 2020), intentando atribuirle características morales humanas como la idea de cultivar una «IA confiable», o el exceso de confianza que se atribuye al ChatGPT como autoridad epistémica, donde se asume una única verdad sin una base adecuada en la evidencia (Cooper, 2023).

En este plano de las sombras, se amplía la mirada a impactos sistémicos (como los ambientales, laborales, de poder o gobernanza). Nos referimos también a riesgos de la IA generativa como los que suponen entrenar y usar modelos que consumen mucha energía y agua y aceleran la renovación de hardware.

Igualmente, se destaca que parte del trabajo necesario para su desarrollo (como las tareas de anotación o codificación de contenidos que permitan el entrenamiento de modelos) recae en personas con condiciones precarias. El control de datos, modelos y nubes se concentra, por otra parte, en pocas empresas, creando grandes dependencias y dificultando la gobernanza. Igualmente, hay problemas de seguridad y fiabilidad (como las manipulaciones de prompts o las respuestas inventadas o «alucinaciones») y sesgos que afectan más a lenguas y grupos poco representados, entre otros. Todos estos elementos invitan, en los próximos años, al refuerzo de la supervisión humana significativa, la transparencia y la introducción de mecanismos de gobernanza antes de aplicar estas herramientas en ámbitos sensibles.

Sin dejar de considerar estos u otros elementos críticos, como algunos de los citados de manera escueta anteriormente, si pensamos en las avenidas que se ofrecen actualmente a las ciencias sociales al amparo de los avances más recientes en la inteligencia artificial, algunas que nos parecen notorias para las ciencias sociales son:

1. El mayor alcance y escala con el que podemos desarrollar nuestra investigación, que nos permiten analizar un volumen y variedad de datos antes impensable, incluyendo datos de diferente tipo (estructuras y no estructurados) y procedentes de diversidad de fuentes;
2. la identificación de patrones, correlaciones, estructuras, redes o tendencias que no eran palpables con métodos tradicionales;
3. la posibilidad de mejorar nuestro conocimiento y articulación de las dimensiones macro, meso y microsociales, así como poder llevar a cabo una combinación de resultados de investigación con mayor precisión al contar con herramientas analíticas más potentes, con capacidad ampliada para recopilar y procesar datos, sean estos «small» o «big data»;
4. la oportunidad de generación de nuevas hipótesis y líneas de trabajo, ante los nuevos hallazgos que se van obteniendo;
5. la versatilidad de las nuevas herramientas, adaptables y con posibilidades de personalización a diferentes sectores sociales o personas;
6. la mayor capacidad de visualizar datos complejos, de forma interactiva, permitiendo una mejor comprensión de la realidad;
7. la mayor facilidad para la obtención de evidencias sólidas para una mejor toma de decisiones, así como para fundamentar una variedad de políticas o programas sociales;
8. la potenciación del desarrollo de métodos mixtos, inter y transdisciplinares para el conocimiento de la realidad social, facilitando la colaboración entre científicos y expertos de los diferentes dominios científicos (Ciencias Sociales, Artes y Humanidades, Ciencias de la Vida y Biomedicina, Ciencias Físicas y Tecnología);
9. la automatización de tareas pesadas o que consumen mucho tiempo en la investigación, ganando en eficiencia;
10. la posibilidad de reforzar en las ciencias sociales las aproximaciones basadas en métodos mixtos y las estrategias de combinación metodológica.

Una de las ganancias claves con el desarrollo de la inteligencia artificial y su implementación en las ciencias sociales es la ampliación de nuestra capacidad para comprender fenómenos complejos, observar y adelantarnos a tendencias y la formulación de intervenciones sociales más efectivas y basadas en evidencia. Un aspecto, por otra parte, que nos parece esencial y sobre lo que ya han venido advirtiendo algunas investigaciones recientes en diferentes áreas es avanzar en la implementación de la colaboración humano-IA, así como en la necesidad de la alfabetización en la IA, más que confiar críticamente en estas tecnologías (Harrer, 2023; Walter, 2024).

Concluyendo, desde nuestra perspectiva, uno de los aspectos en lo que las ciencias sociales pueden obtener más ganancias al apostar por estos campos en desarrollo es precisamente en potenciar el alcance de los métodos mixtos o híbridos, y de las líneas de trabajo cualitativas que, en cierto modo, quedaban oscurecidas con la eclosión de los *big data* y el énfasis en la detección de patrones a escala macro. El trabajo de Gómez Espino (2026) en esta sección de DEBATE, que pone el acento en cómo aplicar la IA y los LLM para mejorar nuestra capacidad de desarrollar transcripciones y codificaciones cualitativas con la herramienta *Social Verbatim*, es un buen ejemplo de ello, que, a su vez, nos hace pensar en la artificialidad de los límites entre lo cuantitativo y lo cualitativo, ayudando a diluirlos. La posibilidad de implementar modelos de IA tipo *zero-shot*, como explica en su artículo, que puedan ser entrenados por científicos sociales para identificar o categorizar temas emergentes o «códigos», es uno de los desafíos que se sugieren. Igualmente, este trabajo, así como el de Gendler (2026), nos remite a la importancia que tienen en las ciencias sociales una aproximación ética y no crítica a los datos por los efectos sociales y sensibilidad que ello comporta.

4. Financiación

Este artículo se beneficia de nuestra implicación en el Proyecto I+D+i titulado «Teorías de la conspiración y discurso de odio online: Comparación de pautas en las narrativas y redes sociales sobre COVID-19, inmigrantes, refugiados y personas LGBTI [NON-CONSPIRA-HATE!]», PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER/EU, siendo uno de los aspectos claves en este proyecto el estudio de los impactos y aplicaciones de la inteligencia artificial.

5. Bibliografía

- Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10), 6187-6203. <https://doi.org/10.1080/10494820.2023.2253861>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E. y Herrera, F. (2011). An approach to detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, 5(1), 146-166. <https://doi.org/10.1016/j.joi.2010.10.002>

- Codd, E. F. (1970). A relational model of data for large shared data banks. *Commun. ACM* 13, 6 (junio), 377–387. <https://doi.org/10.1145/362384.362685>
- Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, 32, 444–452. <https://doi.org/10.1007/s10956-023-10039-y>
- Gendler, M. A. (2026). Ciencias sociales y tecnologías digitales: un largo y complejo camino de enfoques e interrelaciones. *Revista Centra de Ciencias Sociales*, 5(1), 171–192. <https://doi.org/10.54790/rccs.175>
- Gómez Espino, J. M. (2026). Los LLM y la codificación en la investigación cualitativa: avances y oportunidades para Social Verbatim como herramienta integral cualitativa. *Revista Centra de Ciencias Sociales*, 5(1), 193–216. <https://doi.org/10.54790/rccs.176>
- Gualda, E. (2025). Inteligencia artificial generativa, grandes modelos de lenguaje (LLMs) y analítica aumentada vs. big data y ciencia de datos: Nuevas avenidas para la investigación social — Dataset (XLSX), Script (R) y Visualización HTML (Plotly) — Materiales complementarios [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.17298490>
- Harrer, S. (2023). Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *EBioMedicine*, 90, 104512. <https://doi.org/10.1016/j.ebiom.2023.104512>
- Luhn, H. P. (1958). A Business Intelligence System. *IBM Journal of Research and Development*, 2, 4, 314–319. <https://doi.org/10.1147/rd.24.0314>
- Meskó, B. (2023). Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. *Journal of Medical Internet Research*, 4, 25, e50638. <https://doi.org/10.2196/50638>
- Mills, C. W. (1959). *La imaginación sociológica*. México: Fondo de Cultura Económica.
- Mügge, D. (2024). EU AI sovereignty: for whom, to what end, and to whose benefit? *Journal of European Public Policy*, 31(8), 2200–2225. <https://doi.org/10.1080/13501763.2024.2318475>
- Murdoch, B. (2021). Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics*, 22, 122. <https://doi.org/10.1186/s12910-021-00687-3>
- Pinto-Coelho, L. (2023). How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. *Bioengineering*, 10(12), 1435. <https://doi.org/10.3390/bioengineering10121435>
- Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. *Sci Eng Ethics*, 26, 2749–2767. <https://doi.org/10.1007/s11948-020-00228-y>
- Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society*, 75, 102372. <https://doi.org/10.1016/j.techsoc.2023.102372>

- Skinner, R. E. (2012). *Building the Second Mind: 1956 and the Origins of Artificial Intelligence Computing*. Smashwords. UC Berkeley. <https://escholarship.org/uc/item/88q1j6z3>
- Van Eck, N. J. y Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, 3, 1, 100165. <https://doi.org/10.1016/j.jjimei.2023.100165>
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. *Int J Educ Technol High Educ*, 21, 15. <https://doi.org/10.1186/s41239-024-00448-3>

Estrella Gualda

Catedrática de Sociología en la Universidad de Huelva, académica de número de la Academia Iberoamericana de La Rábida y directora del grupo de investigación «Estudios Sociales e Intervención Social» (ESEIS). En los últimos años ha prestado gran atención a la sociología computacional, ciencia de datos, *big data*, análisis de redes sociales, así como a cuestiones tales como las teorías de la conspiración, los discursos de odio en línea y la desinformación en relación con la COVID-19, y las personas inmigrantes, refugiadas y LGTBIQ+, encontrándose sus trabajos en revistas y editoriales de gran prestigio como *Nature Communication*, *Nature Human Behaviour*, *Nature Scientific Data*, *IEEE Access*, *Array*, *Frontiers in Psychology*, *PNAS Nexus*, *Political Psychology*, *The American Sociologist*, *REIS*, *Empiria*, *Gazeta de Antropología*, *Redes*, *Springer*, *Routledge*, *Dykinson*, entre otras.

Notas

1 Nos referimos a términos incluidos en el mapa o procedentes de la base de datos completa, para una mejor descripción.