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

Editorial: Generative Artificial Intelligence, Large Language Models (LLMs) and Augmented Analytics vs Big Data and Data Science: New Avenues for Social Research

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

Estrella Gualda
University of Huelva (UHU, ESEIS/COIDESO/CISCOA-Lab), Spain
estrella@uhu.es

Received/Recibido: 9-6-2025
Accepted/Aceptado: 14-10-2025

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 AI; augmented analytics; Web of Science; knowledge maps; scientific publications; social sciences.

HOW TO REFERENCE: 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

The English version can be read at https://doi.org/10.54790/rccs.150

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 «candentes» 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 artificial 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 artificial generativa; analítica aumentada; Web of Science; mapas de conocimiento; publicaciones científicas; ciencias sociales.

1. From big data and data science to generative artificial intelligence, large language models and augmented analytics

Although we may be at the outset of an emerging trajectory whose implications for the social sciences are still unclear, substantial transformations are already anticipated in several key aspects of intellectual work and research approaches. These changes stem from significant advances associated with technologies and tools such as large language models (LLMs), generative artificial intelligence and augmented analytics. Recent innovations are diversifying and amplifying the impacts that artificial intelligence, business intelligence, big data and data science were already having on the social sciences.

The social impact and the transformations resulting from these innovations – particularly those driven by generative AI – are already under way (Saetra, 2023), affecting multiple sectors and domains such as education (Walter, 2024; Chiu, 2024) and healthcare, where some speak of a paradigm shift in the field of medical imaging (Pinto-Coelho, 2023).

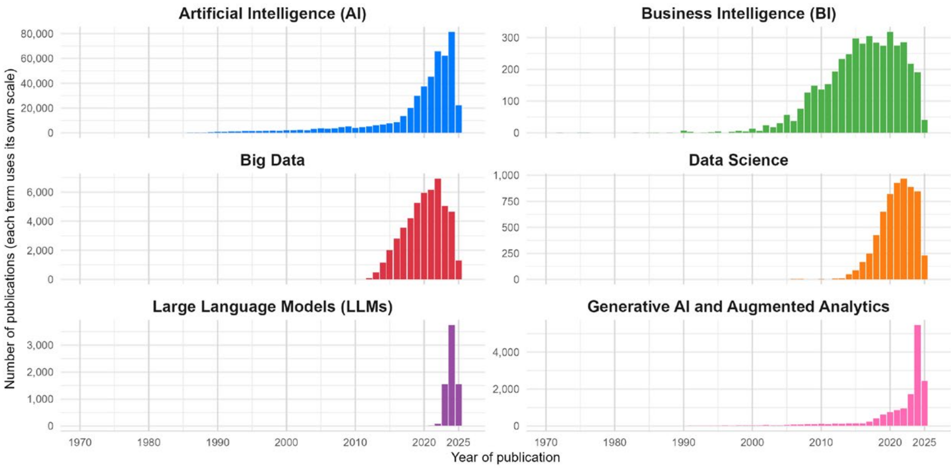
Based on our understanding of the social sciences – as well as insights drawn from texts by Gendler (2026) and Gómez Espino (2026) included in this Discussion section of the *CENTRA Journal of Social Sciences* (https://www.centracs.es/revista) – we anticipate that the coming years will witness the emergence of research strategies that incorporate one or more of the advances mentioned above. These strategies will likely be combined with varying degrees of engagement in prompting or prompt engineering (Walter, 2024), which is already enabling the social sciences (as well as other disciplines such as health sciences – see Meskó, 2023) to benefit from these innovations without the previous constraint of requiring advanced programming skills. This surge in research that leverages such technological developments is not incompatible with more classical or hybrid methodological approaches.

The trajectory of scientific research in these fields, technologies and tools becomes evident through comparative analysis of relevant publications – a task we summarise below due to page constraints. To provide an introductory comparative overview, a series of bar charts was generated using R. These charts illustrate the temporal evolution of publications whose titles or abstracts include specific terms, based on six separate bibliographic searches: (1) "artificial intelligence OR AI"; (2) "business intelligence"; (3) "big data"; (4) "data science"; and (5) "large language models OR LLMs". Lastly, in order to identify more emerging areas linked to developments in analytical artificial intelligence and augmented analytics, a sixth search was conducted using key terms such as (6) "generative AI", "generative artificial intelligence", "diffusion models", "GANs", "augmented analytics", "automated insights", "natural language generation" and "AI-driven analytics". These general and specific terms offer an overview of ongoing research in this area and enable a preliminary assessment of how these developments may contribute to enriching the work of the social sciences. The searches were carried out using the Web of Science (WoS) bibliographic database (https://www.webofscience.com/wos/), covering the entire series of available publications up to the end of May 2025. Publications were filtered by the "Social Sciences" research domain in WoS.

Figure 1 presents absolute annual publication counts for each search. To improve the readability of temporal trends, each panel employs an independent vertical axis (scaled to the range of its respective series). It is important to note that bar heights are not comparable across panels, only within individual panels. Given the differences in volume between areas, we opted for this panel design with separate axes in order to clearly show the evolution in each case, without resorting to logarithmic scales. No normalisation was applied to the series: the bars represent absolute counts per year within each panel.

Figure 1

Trends in publications in the social sciences (by field and technology) – WoS



Source: compiled by the author using R, based on data from WoS.
An interactive version of this chart is available. Both the interactive chart and the code and data used to generate it can be accessed via the following link: https://doi.org/10.5281/zenodo.17298490 (Gualda, 2025).

This comparative strategy provides a valuable overview of the evolution of these fields and highlights key milestones. It also lays the groundwork for a more in-depth exploration of the core terms used to define these fields and technologies, and the connections between them. In addition, it offers insight into some of their most significant branches through network or co-occurrence analysis of key concepts (knowledge maps), a method previously applied in other bibliometric studies (Cobo *et al.*, 2011).

Although all available publication data from WoS were extracted without a time restriction, following a preliminary pre-analysis, the comparative time series visualised in the charts focuses on the period from 1970 to the end of May 2025 (the cut-off point at the time of writing). In the period preceding the 1970s – particularly from the 1950s onward – the field of artificial intelligence was already taking shape. From the outset, it was grounded in the notion that machines could be created to simulate human thought or to perform tasks that "would be considered intelligent if done by a human being" (Skinner, 2012, p. 3). Even within the social sciences, early explorations can be traced concerning the use of computers for statistical analysis and the simulation of social phenomena, notably with the emergence of landmark software such as SPSS in 1968. Gendler's (2026) contribution to this Discussion section offers a detailed analysis of this historical evolution in the social sciences and humanities in relation to digital technologies.

Meanwhile, the early development of business intelligence can be traced back to the 1960s (Luhn, 1958), with the introduction of foundational concepts and systems for organising and analysing business data to support decision-making. Similarly, the creation of relational databases laid the foundations for data organisation (Codd, 1970), serving as precursors to modern data management systems.

A comparison of the publication trends across the six identified fields and technologies shows that academic work on artificial intelligence related to the social sciences began significantly earlier than in the other areas. A substantial increase is visible from around 2010 onwards, with particularly rapid growth over the past five years (see Figure 1). Interest in artificial intelligence in the social sciences has remained steady over time, although its growth was initially slow during the early decades. By contrast, while publications on business intelligence appeared slightly later than those on artificial intelligence, they began to increase steadily from the early 1990s. From 2005 onwards, a sustained and marked growth in publication volume can be observed. It should be emphasised that the trends presented here refer specifically to the evolution of literature on these fields and technologies, filtered by publications classified under the "Social Sciences" domain in the Web of Science database.

Compared to the earlier emergence of the fields of artificial intelligence and business intelligence, the domains of big data and data science – which are often closely linked – began to attract growing interest within the social sciences from the second decade of the 21st century. A surge in research output relating to big data is evident from around 2010–2012, a period that also saw a parallel, albeit somewhat more modest, increase in publications associated with data science, a field closely connected to big data. In the past five years, there appears to have been a relative decline in the number of documents referring specifically to big data, contrasted with a continued rise in publications focused on data science.

Key developments shaping this evolution in the 21st century include the advances in and expansion of deep learning (DL) since 2010, which have revolutionised the field of AI, and the emergence of large language models (LLMs), somewhat later, with the introduction of technologies such as ChatGPT and similar tools (Meskó, 2023; Saetra, 2023; Cooper, 2023). These advances, particularly in deep learning, have enabled progress in areas such as image and voice recognition and natural language processing – developments that are having a profound impact on the social sciences. At the same time, the creation of language models based on transformer architectures, capable of generating and interpreting text at great speed, constitutes a new and significant disruption within the field. This topic is addressed in detail in Gómez Espino's (2026) article in this Discussion section, which focuses on thematic qualitative coding processes. It examines the contributions of large language models to qualitative research through the development of the Social Verbatim tool, emphasising the importance of human–machine collaboration.

Although publications on large language models, generative intelligence and augmented analytics represent very recent areas of research, the exponential increase in related outputs – visible in the charts in Figure 1 – suggests that these domains will undergo intense development within the social sciences in the coming years. LLMs and the practice known as prompting may provide a significant boost to a wide range of tasks associated with research in general, and with social science research in particular, by stimulating what Mills (1959) referred to as the "sociological imagination". The potential of artificial intelligence to act as a technical assistant – in its latest forms (LLMs, generative and analytical AI) – may help the social sciences to overcome previously existing technical barriers. It can also enhance research across a wide spectrum of activities, including literature retrieval, document coding and classification, text mining, rapid execution of complex analyses across diverse data types (text, image, sound, video), improved visualisations and much more.

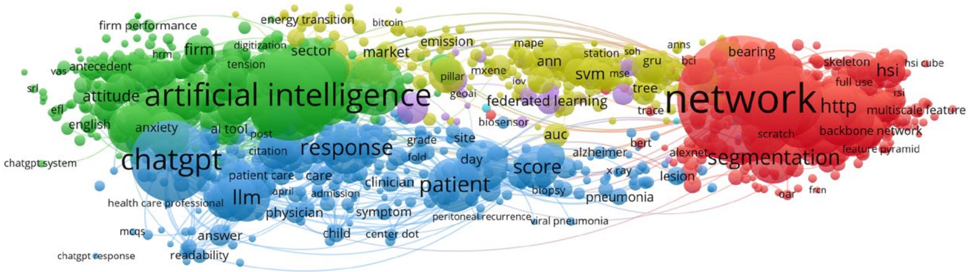
2. Knowledge maps and significant clusters related to literature on classical and modern artificial intelligence, business intelligence, big data and data science

Having broadly outlined the evolution of publications in the social sciences across different fields and technologies, we now turn to an analysis of co-occurrence networks and term clusters, allowing us to sketch out knowledge maps (Cobo *et al.*, 2011) of the field of artificial intelligence – both classical and modern – as well as related lines of research such as business intelligence, big data and data science. Due to space constraints, this will be a concise global analysis, although some additional details may be observed in the accompanying figures. For this analysis, we focus exclusively on the "most cited" and "hot" documents indexed in WoS, in order to visualise both the key topics and concepts emerging from this body of literature and the interconnections among them.

To this end, from the set of articles used to trace publication trends, we extracted those classified in WoS as "most cited" and "hot" across the six search categories (artificial intelligence, business intelligence, big data, data science, large language models, and generative and augmented AI). These datasets were then merged using R, duplicate records were removed and the resulting file was adapted for import into the bibliometric analysis software VOSviewer (Van Eck and Waltman, 2010). Relevance metrics were calculated, followed by the visualisation and mapping of the most significant clusters.

Figure 2 presents a knowledge map that displays the most relevant terms appearing in the most cited works in this literature. Network analysis conducted with VOSviewer enabled the identification of multiple clusters, each grouping together interconnected terms that represent core thematic areas within the field. Terms were extracted from the titles and abstracts of the documents. The visualisation includes only highly relevant terms, which provide clear indications – through colour-coded subgraphs – of the principal research areas. The interconnections allow us to identify the most prominent links within the literature. In the brief analysis that follows, we refer to some of the most influential terms based on the Total Link Strength indicator – a metric that highlights the foundational elements of each cluster. The size of each node within the cluster is also associated with the centrality of that term. The proximity between terms likewise suggests a stronger relationship or frequency of co-occurrence.

Figure 2
Knowledge map of fields and technologies associated with artificial intelligence



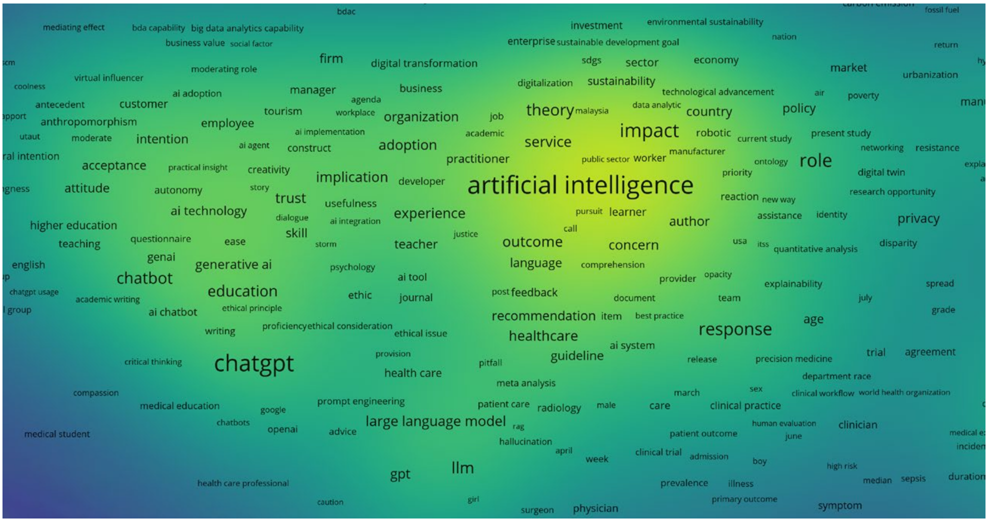
Source: own research based on VOSviewer. Data extracted from WoS and processed in R.

2.1. Artificial intelligence, society and human–machine interaction (green cluster)

The green cluster, focused on artificial intelligence, constitutes one of the core groupings in this knowledge map. It reveals the centrality of literature addressing the adoption, perception, use and societal impact of AI and its related technologies across a range of social and organisational contexts. This includes domains such as customer experience, business management and user acceptance. Notably, some of the most prominent terms[1] include “adoption”, “acceptance”, “AI adoption”, “AI technology”, “AI use”, “usage intention” and “technology acceptance model (TAM)”. These are linked to studies examining the psychological and behavioural factors that influence technology adoption – particularly perceived usefulness and ease of use. Many of these and related terms are visible in the heat map shown in Figure 3.

Figure 3

Heat map or density map of the green cluster



Source: own research based on VOSviewer. Yellow indicates a higher density or concentration of publications associated with the given topic.

This cluster also includes a substantial body of literature on human–AI and human–robot interaction, illustrated by terms such as "human AI interaction", "human robot interaction (HRI)", "human machine interaction" and "robot". Complementary lines of research address human emotions and attitudes – terms such as "anxiety", "AI anxiety", "fear", "negative emotion" and "positive emotion" feature prominently – as well as robot characteristics that influence such interactions. These include "anthropomorphism" and "social presence", which are frequently cited in the literature.

The cluster also places strong emphasis on the role of AI in education and learning, with key terms including "education", "higher education", "student", "teacher", "academic writing" and "language learning". The literature explores how AI is reshaping teaching and learning processes, while also engaging with ethical challenges such as plagiarism, privacy and broader ethical considerations. In the business domain, the literature documents the impact of AI on management and digital transformation, highlighting its role in redefining the organisational landscape. This is reflected in the presence of terms such as "business", "digital transformation", "innovation", "sustainability", "job", "employment", "supply chain" and "customer", pointing to its relevance across a variety of sectors including hospitality, tourism and services. While not the only cluster addressing these developments, this grouping is particularly notable for its focus on LLMs and generative AI. This is evident in the frequent use of terms such as "chatbot" and "ChatGPT", which reflect ongoing experimentation in these areas and contrast with earlier research centred primarily on predictive models. The cluster also reveals a strong orientation towards rigorous research and methodological robustness. Terms such as "empirical study", "qualitative study" and "SEM" (structural equation modelling) are common, underscoring the importance of empirical validation and theoretical development in advancing our understanding and management of AI's societal impacts.

2.2. Deep neural networks: computer vision, detection and image segmentation (red cluster)

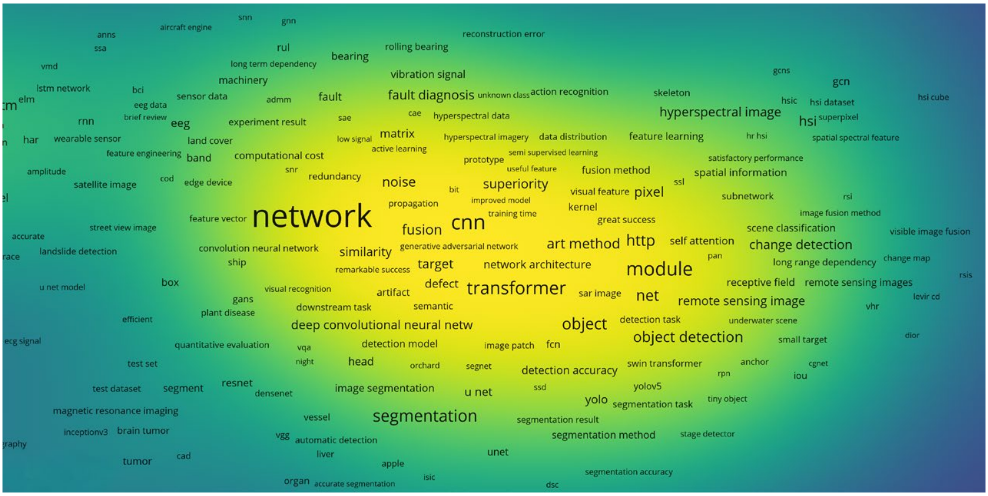
This cluster encompasses core concepts related to architectures such as convolutional neural networks (CNNs) and transformers. The associated literature highlights key applications in recent artificial intelligence research – applications that are of growing relevance and potential within the social sciences – particularly in the areas of computer vision, object detection and image segmentation.

The publications grouped within this cluster represent a focal point of current AI research, notable for their emphasis on the potential of image processing and analysis, among other areas. The term "network" exhibits the highest degree of connectivity within the cluster. Other key terms around which the literature is organised (Figure 4) include: "module", "segmentation", "detection", "fault diagnosis", "target detection", "neural", "image fusion", "layer", "experimental result", "code", "github.com" and "superiority". Algorithm-related terms such as "CNN", "LSTM" (long short-term memory) and "SVM" (support vector machine) also occupy a central position.

The literature suggests that research in this area is primarily focused on the design of efficient architectures, the ongoing improvement of performance and the applicability of models to a variety of real-world scenarios. This includes efforts in feature extraction and fusion, as well as the development of models that are efficient and deliver excellent performance. Publications in this cluster are also characterised by a strong commitment to experimentation, empirical validation and collaborative practices, particularly through code sharing. These studies seek to address challenges such as class imbalance and overfitting, with the aim of enhancing generalisation ability.

Figure 4

Heat map or density map of the red cluster



Source: own research based on VOSviewer. Yellow indicates a higher density or concentration of publications associated with the given topic.

2.3. AI and LLMs in the healthcare domain: from diagnosis to decision-making and ethical considerations (blue cluster)

This cluster is large and heterogeneous, bringing together research in artificial intelligence (AI) with the healthcare and medical domains. It is primarily oriented towards the application of AI and data science in healthcare, with particular emphasis on the growing use of LLMs such as ChatGPT. Key terms associated with this cluster include "ChatGPT", "LLM", "healthcare", "medicine", "patient", "clinical", "diagnosis", "disease", "symptom", "physician", "clinical decision", "drug discovery", "pneumonia" and "COVID".

Research in this area focuses on the diagnosis and detection of diseases – such as cancer, pneumonia and Alzheimer's disease – using AI to enhance accuracy and efficiency. These efforts often rely on medical imaging techniques, including chest X-rays and computed tomography (CT) scans. The literature also examines the role of AI in supporting clinical decision-making processes and in the management and analysis of patient data. In parallel, this cluster engages extensively with the ethical, social and regulatory implications of AI in healthcare. It addresses principles such as accountability, transparency and fairness, while also confronting challenges related to algorithmic bias and misinformation. Research methodology is a central concern in this body of work, highlighting the need for a robust evidence base – grounded in clinical trials and systematic reviews – to ensure the safe, effective and equitable integration of AI into medical practice. Overall, this cluster reflects the comprehensive ecosystem of AI in the healthcare sector, encompassing both its technological capabilities (diagnosis, clinical support, data processing) and the critical ethical and social considerations that shape its implementation.

2.4. Distributed AI, blockchain, privacy and optimisation (purple cluster)

This cluster represents the intersection of artificial intelligence with distributed technologies, security, privacy and optimisation. The associated terms indicate a focus on the secure and efficient implementation of AI within distributed or decentralised systems, with an emphasis on confidentiality, integrity and scalability in large-scale applications. Key terms include "blockchain technology", "distributed ledger technology (DLT)", "federated learning", "edge computing", "NFT" (non-fungible token) and "cryptocurrency" in the field of distributed technologies. Security- and privacy-related concepts are also central to this cluster, as shown by the presence of terms such as "data privacy", "privacy policy", "privacy preserving", "privacy leakage", "privacy protection", "attack", "cybersecurity", "distributed attack" and "malicious attack".

In the area of optimisation and operational efficiency, terms such as "optimization", "algorithm", "resource allocation" and "energy efficiency" feature prominently. Emerging application areas identified in this cluster include "metaverse", "digital twin", "smart city", "smart manufacturing", "internet of vehicles" and "cyber physical system". Other key terms include "content generation", "recommender system", "deep reinforcement learning", "graph neural network", "data management", "data sharing" and "simulation". Research in this area focuses on the secure and efficient operation of AI-driven systems in contexts such as federated data environments, supply chain management, smart infrastructure and the metaverse. Particular emphasis is placed on information protection and the prevention of cyberattacks, with simulation playing a central role as a methodological approach for validating proposed solutions.

2.5. Machine learning, predictive analytics and business intelligence (yellow cluster)

This cluster maps concepts related to advanced machine learning and deep learning techniques and models applied to data analysis and predictive tasks. It includes architectures such as "LSTM", "SVM", "random forest", "XGBoost" and "ANN" (artificial neural network). Key terms in the associated literature also reference validation metrics, including "AUC" (area under the curve) and "RMSE" (root mean square error), which are commonly used to assess model performance. The cluster additionally encompasses "business intelligence" (BI) and the application of these models to business and market contexts, as reflected in terms such as "market" and "firms".

At a broader level, the different clusters in the map exhibit strong interconnectivity. The green cluster (artificial intelligence), in particular, appears to serve as a central axis, linking directly with the blue cluster (AI and LLMs in healthcare), the yellow cluster (machine learning and BI), and the red cluster, which focuses on recent developments in deep neural networks and computer vision. This structural configuration underscores the role of AI as a unifying core that connects both technical advancements and applied domains.

3. Avenues for social research: by way of conclusion

Although the brief and general overview we have provided – drawing from a single source (Web of Science) – can offer only a preliminary outline of the profound transformations currently under way, it is evident that the new technologies and tools associated with artificial intelligence are opening up an immense range of opportunities for research, analysis and intervention across the diverse disciplines that make up the social sciences. While synthesising this evolving landscape was by no means a simple task, the exercise has served to introduce and frame this Discussion section of the *CENTRA Journal of Social Sciences*, and to provide a preliminary insight into some of the themes that have attracted the most attention within the social sciences at various historical junctures, as well as to highlight emerging areas within AI that deserve particular focus. A more comprehensive analysis is offered in the contribution by Gendler (2026) included in this issue.

Within the broad field of AI, special attention is warranted for developments related to large language models (LLMs) and to generative and analytical AI. These models – often multimodal in nature (Meskó, 2023) – are distinguished by their capacity not only to analyse or classify data, but also to generate original and novel content. The ability to create text, images, audio, video or code represents a highly advanced capability that must not be overlooked.

This generative capacity, alongside its benefits, has also sparked major questions, debates and critiques – especially concerning the potential harms that may arise when generative AI is used to produce content (Saetra, 2023; Harrer, 2023). Prominent concerns include the exponential amplification of disinformation, hate speech, deepfakes, and the creation of harmful or misleading content through unsupervised generative AI systems (Harrer, 2023).

Other critical dimensions relate to the presence of biases embedded in algorithms (Varsha, 2023); accusations of plagiarism involving academic, artistic or other forms of content; ethical considerations; and potential infringements of copyright. These issues intersect with broader challenges concerning data privacy, anonymisation and security, as well as the protection of data from monopolisation by private actors, and the need for robust frameworks of AI ethics and governance (Murdoch, 2021; Mügge, 2024). Further substantive critiques – previously noted in relation to the green cluster – include the tendency to anthropomorphise AI (Ryan, 2020), such as the attribution of human moral qualities to artificial systems through concepts like "trustworthy AI", and the growing over-reliance on tools like ChatGPT as epistemic authorities. This latter concern relates to the uncritical acceptance of outputs as singular truths, often lacking adequate evidentiary foundations (Cooper, 2023).

In this shadow landscape, the focus broadens to consider systemic impacts – including environmental, labour-related, power and governance issues. Among the risks associated with generative AI is the high resource consumption involved in training and deploying large-scale models, particularly in terms of energy and water usage, as well as the accelerated turnover of hardware infrastructure.

Moreover, part of the labour underpinning these systems – such as annotation and content labelling tasks required for model training – is frequently outsourced to individuals working under precarious conditions. Control over data, models and cloud infrastructures is also becoming increasingly concentrated in the hands of a few companies, creating significant dependencies and limiting prospects for democratic governance. Additional challenges include security and reliability concerns (e.g. prompt manipulation or fabricated outputs, often referred to as hallucinations) and biases that disproportionately affect under-represented languages and social groups. Collectively, these concerns call for the reinforcement of meaningful human oversight, greater transparency and the establishment of governance mechanisms before such tools are deployed in sensitive domains.

Without losing sight of these or other critical dimensions – some of which have been noted only briefly – it is also important to reflect on the new avenues currently opening up for the social sciences in light of recent advances in artificial intelligence. Several of these appear especially significant:

1. The expanded scope and scale at which research can now be undertaken, enabling the analysis of previously unimaginable volumes and types of data – including both structured and unstructured data from diverse sources;
2. The identification of patterns, correlations, structures, networks and trends that were difficult or impossible to detect using traditional methods;
3. Enhanced capacity to understand and integrate macro-, meso- and micro-level dimensions of social life, and to combine research findings with greater precision, thanks to increasingly powerful analytical tools and improved data processing capabilities – applicable to both small and big data;
4. The generation of new hypotheses and research pathways arising from previously inaccessible findings;
5. The versatility of emerging tools, which can be adapted and customised for different social sectors, contexts and populations;
6. Improved visualisation of complex data in interactive formats, fostering clearer and more intuitive understandings of social phenomena;
7. Easier generation of robust evidence to support more informed decision-making and to underpin a range of social policies and programmes;
8. Strengthening of mixed-methods, interdisciplinary and transdisciplinary research approaches, facilitating collaboration across scientific domains – including the social sciences, arts and humanities, life sciences and biomedicine, physical sciences and technology;
9. The automation of time-intensive or repetitive research tasks, resulting in greater efficiency and productivity;
10. The reinforcement of methodological integration and the expansion of mixed-methods strategies in social science research.

One of the key benefits associated with the development of artificial intelligence and its implementation in the social sciences is the expansion of our capacity to understand complex phenomena, to observe and anticipate emerging trends, and to formulate more effective, evidence-based social interventions. An additional and increasingly critical aspect – already highlighted by recent research across multiple disciplines – is the need to advance human–AI collaboration, alongside a growing call for AI literacy. Rather than placing blind trust in these technologies, a more reflective and informed engagement is required (Harrer, 2023; Walter, 2024).

In conclusion, from our perspective, one of the most valuable opportunities for the social sciences in engaging with these emerging fields lies in the potential to enhance the reach and depth of mixed or hybrid methods, and to revitalise qualitative research approaches which, to some extent, have been marginalised by the rise of big data and the dominant focus on large-scale pattern detection. The contribution by Gómez Espino (2026), included in this Discussion section, offers a compelling illustration of this potential. His work demonstrates how AI and LLMs can be applied to improve qualitative transcription and coding processes through the Social Verbatim tool. At the same time, it invites reflection on the artificial divide between quantitative and qualitative methods – helping to blur and ultimately transcend this boundary. As he explains, the possibility of implementing zero-shot AI models – capable of being trained by social scientists to identify or categorise emergent topics or "codes" – represents one of the key challenges and opportunities on the horizon. Moreover, this contribution, together with that of Gendler (2026), highlights the importance of maintaining a critical and ethically grounded approach to data within the social sciences, given the social sensitivity and potential implications associated with its use.

4. Funding

This article has benefited from our involvement in the R&D+i project titled "Conspiracy Theories and Online Hate Speech: A Comparative Analysis of Narrative Patterns and Social Networks Related to COVID-19, Migrants, Refugees and LGBTI Individuals [NON-CONSPIRA-HATE!]", PID2021-123983OB-I00, funded by MCIN/AEI/10.13039/501100011033/ and by FEDER/EU. A key component of this project involves the study of the impacts and applications of artificial intelligence.

5. References

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 Medial 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

Estrella Gualda is a full professor of sociology at the University of Huelva, full member of the Academia Iberoamericana de La Rábida and director of the research group Social Studies and Social Intervention (ESEIS). In recent years, her research has focused on computational sociology, data science, big data and social network analysis, as well as on conspiracy theories, online hate speech and disinformation related to COVID-19, and issues concerning migrants, refugees and LGBTIQ+ individuals. Her work has been published in leading journals and by reputable academic publishers, including *Nature Communications, Nature Human Behaviour, Scientific Data, IEEE Access, Array, Frontiers in Psychology, PNAS Nexus, Political Psychology, The American Sociologist, REIS, Empiria, Gazeta de Antropología, Redes, Springer, Routledge* and *Dykinson,* among others.

Notes

1 The terms referenced correspond to those included in the knowledge map or drawn from the full database, to support more precise descriptions.