SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 233

ARTIFICIAL INTELLIGENCE AND ITS ROLE IN THE DEVELOPMENT OF SOCIAL SCIENCES AND HUMANITIES

Zbigniew ORBIK

Silesian University of Technology, Faculty of Organization and Management;
zbigniew.orbik@polsl.pl, ORCID: 0000-0001-7092-1567

Purpose: The purpose of this article is to demonstrate the importance of AI-based tools for advancing research in the social sciences and humanities. The author focuses primarily on how AI can enhance research in these disciplines.
Design/methodology/approach: This article analyses selected available literature on the use of AI-based tools in research conducted in the social sciences and humanities. In addition to a literature review, the author employs the method of analysis and logical construction.
Findings: Artificial intelligence is a tool that representatives of social sciences and humanities have great hopes for. AI raises many important issues for the humanities and social sciences. The emergence of AI technology has changed not only research conducted in the social sciences and humanities but also the paradigm regarding various aspects of practicing the humanities, such as academic writing, learning and intellectual property.
Research limitations/implications: The article refers to a limited number of studies. It discusses the current state of AI's impact on the social sciences and humanities, both in terms of ongoing research and emerging ethical issues.
Originality/value: The article addresses the relatively rarely analysed issue of the impact of AI-based tools on research in the social sciences and humanities. The incredibly dynamic development of AI creates new, previously unknown possibilities, but also raises a number of new problems and challenges.
Keywords: AI, humanities, social sciences.
Category of the paper: Conceptual paper.

1. Introduction

Artificial Intelligence (AI) is rapidly moving beyond speculation to practice across academic disciplines, including the social sciences and humanities. AI's integration into these fields is shifting not only how research is being done but also how knowledge is being constructed and advanced (Hutson, 2024). To understand the impact of AI technologies on methodologies, data analysis, and theory in the social sciences and humanities, this paper

specifically addresses the following research question: How does the incorporation of AI technologies influence the methodologies, data analysis processes, and theoretical developments within the social sciences and humanities?

The social sciences and humanities are based on the study of human societies, behaviours, and cultures, with qualitative and interpretative research techniques dominating these fields. AI has automated traditional methods such as transcription and coding, made large-scale textual and visual data analysis possible, and supported the study of ancient manuscripts and global literature using NLP and machine learning. It allows for the identification of societal patterns and simulation of alternative historical scenarios, thus expanding the capacity for both analysis and hypothesis testing. Therefore, the relevance of this topic is due to the possibility of increasing the efficiency of the research process, but also for providing tools to question and/or deepen our knowledge about human societies.

The aims of this scientific article are threefold: (1) to show how AI is being implemented into methodological development for data processing and analytic techniques, as well as for the automation of traditional research activities, within the social sciences and humanities; (2) to explore the impact of AI technologies on theoretical development within the social sciences and humanities, from hypothesis development to theory generation and interpretation of results; and (3) to identify key ethical and epistemological issues regarding the utilization of AI technologies for social sciences and humanities research, in the form of risks or challenges such as bias, transparency, and accountability.

This paper is based on a thorough review of literature on artificial intelligence, digital humanities, and AI in social sciences research. I consulted empirical studies on AI and AI ethics, analyses of how digital humanities use AI technologies in social science and humanities disciplines, and practical or theoretical frameworks regarding responsible AI for its developers and organizations implementing AI.

AI research is developing rapidly, especially in the social sciences and humanities fields. This field is starting to create practical applications by automating tasks, enabling digital preservation, or fostering collaboration between areas. Meanwhile, several studies showed that AI technology has many potential issues such as bias, data privacy concerns, accountability, fairness, and transparency.

The structure of this scientific paper is as follows: Chapter 2 examines the role of AI in social sciences and humanities research, focusing on applications, methods, and theoretical developments. Chapter 3 addresses the most relevant challenges of AI in the social sciences and humanities research, such as algorithmic bias and data privacy. The last chapter synthesizes and offers reflections regarding this topic.

2. Methods

The article uses a research technique such as literature review to collect, analyze, and synthesize relevant information from a wide range of sources. A literature review is a fully-fledged research method. A well-conducted literature review allows you to get acquainted with the current state of knowledge in a given field and provide reliable answers to research questions (Snyder, 2019). This is an excellent method for analysing interdisciplinary research. It also allows for the synthesis of research results and the discovery of areas where further research is necessary. The literature review was conducted on the Google Scholar and Science Direct platforms. The databases were searched based on the following keywords: sustainable development, progress, environmental ethics, sustainability.

The table 1 below summarizes the different types of literature review.

Table 1.
Types of Literature Review

Type	**Systematic**	**Semi-systematic**	**Integrative**
Typical purpose	Synthesize and compare evidence	Overview research area and track development over time	Critique and synthesize
Search strategy	Systematic	May or may not be systematic	Usually not systematic
Research questions	Specific	Broad	Narrow or broad
Analysis and evaluation	Quantitative	Qualitative/quantitative	Qualitative
Sample characteristics	Quantitative articles	Quantitative articles	Research articles, books, and other published texts
Examples of contribution	Evidence of effect. Inform policy and practice	State of knowledge. Themes in literature. Historical overview. Research agenda. Theoretical model.	Taxonomy or classification

Source: own elaboration based on: Snyder, 2019.

The considerations presented in the article also use the method of logical analysis. The essence of this method is the use of logical tools to distinguish correct arguments, i.e. those that are consistent with the laws of logic, from those that violate the rules. Logic and its methods allow us to discover the most general structural aspects of our thinking and actions (Peregrin, Svoboda, 2017). The application of the tools of logical analysis to propositions expressed in natural language is justified by the fact that the rules of such a language are also subject to the laws of logic.

Logical analysis is helpful in research conducted in social and management sciences. It enables better understanding of their theoretical and conceptual framework (Ngulube et al., 2015). Management sciences, like any other scientific discipline, have their own research methods and are characterized by interdisciplinarity. For this reason, some kind of methodological eclecticism can be attributed to these sciences. It allows to combine methods from various disciplines, paradigms and approaches, ensuring their cognitive and pragmatic

efficiency. These are both quantitative and qualitative methods. The research process in management sciences is often accompanied by various types of models, including: theoretical and conceptual, research or statistical. The model, in general, means a hypothetical thought structure, i.e. the arrangement of assumptions, concepts and relationships between them, allowing you to describe some aspect of reality. The role of logic in research is crucial because it allows to draw correct conclusions and avoid errors in reasoning.

Logic is a formal science. Both formal and informal principles of logic can be applied to the analysis of a text in natural language most often used in social sciences and humanities (Ngalande, 2011). The author of this article does not use formalisms. In the social sciences, universal logical principles are used to determine the correctness of statements and theories formulated within these disciplines (Mahoney, 2021).

3. The Role of AI in Social Sciences and Humanities Research

Artificial intelligence is changing research methods in the social sciences and humanities, providing new options for data analysis, visualization, and the development of theories. In this context, new opportunities and dangers emerge. These include the potential for research bias, ethical challenges, and a need for transparency.

3.1. Current Applications and Methods

AI tools have become increasingly significant in social science and humanities research methods by innovating qualitative and quantitative data collection, processing, and interpretation techniques. AI-based transcription and coding solutions have improved data management by removing human error, thereby facilitating researchers to focus more on analysis. This has aided social science research in terms of efficiency and analysis (Christou, 2023). While the use of these AI tools helps social sciences research by saving time and minimizing mistakes, researchers need to be mindful of the risk of oversimplifying complex social scenarios with these technologies.

Advances in pattern-recognition with the aid of AI tools have helped social science researchers in identifying themes and patterns in qualitative datasets. These AI tools can discover insights beyond human ability in a rapid way (Unke et al., 2021). While this has opened new avenues for analyzing research datasets, researchers need to be aware of the effects of algorithmic processes on interpretations of cultural and contextual issues.

Transcription AI, particularly through speech-recognition technologies, has greatly helped researchers in saving time on data preparation tasks, such as transcriptions of interview or focus group discussions. This allows the fast and reliable generation of transcripts at low cost (Mojadeddi, Rosenberg, 2024). This can allow researchers to conduct research projects by

themselves and as a group. Despite its strengths, this also has challenges in transcribing poor-quality audios and in cross-language transcriptions and needs to be done in a manner that maintains reliability.

AI-based data visualization has helped researchers convey complex qualitative data through techniques such as word clouds or thematic maps. These are very useful in collaborative projects and across different social sciences fields (Arulselvi, 2024). As this method of analysis is emerging, it is important that researchers balance the use of AI in explaining complicated topics with the potential risk of oversimplification or distortion.

LLMs, such as ChatGPT, have been used to generate high-variance survey responses to help social scientists in testing out survey tools without having to depend on real-world respondents and being affected by biases. This approach allows the evaluation of survey reliability and validity (Grossmann et al., 2023). As such, the challenge for this method is finding ways to verify that the answers generated are genuine. Moreover, these can be used to analyze historical data and answer "what-if" questions about alternative historic narratives. By using LLMs, researchers can study social phenomena or test theoretical relationships previously not explored (Grossmann et al., 2023). Thus, one needs to analyze whether the use of these LLMs has an effect on the theoretical frameworks used.

Text analytics and natural language processing have opened pathways for the study of press material, historical literature, and other text datasets that would have taken researchers long periods of time to analyze and search patterns (Chapinal-Heras, Díaz-Sánchez, 2023). This has enabled scholars in the humanities and social sciences to identify concepts within large texts to look for patterns and make comparisons between different texts from different periods. For instance, this has been used to investigate conceptualizations used by ancient Greek philosophers in old literature (Köntges, 2020). These kinds of AI tools have allowed for exploring patterns that, prior to the development of these, were simply beyond the capabilities of researchers to do so. While the application of text analysis has helped the humanities in a multitude of ways, it is important to ensure that there is limited oversight and that these tools are not relied on as a definitive arbiter of textual meaning.

By identifying concepts, themes, and sociological constructs, AI allows researchers in the humanities to make comparative analyses to look for patterns. This helps researchers to understand culture in different scenarios (Ho, Vu, 2025). When used in these circumstances, it assists researchers in the humanities to contextualize multilingual and multi-sourced data, supporting diversity and pluralism in this sector (Chapinal-Heras, Díaz-Sánchez, 2023). Thus, there needs to be a constant check on these AI techniques to see if they are doing accurate comparative assessments to see if various factors influence it.

These AI tools have enabled pattern identification in historical imagery and object databases, which leads to the advancement of conservation analysis in the cultural heritage sector. With AI analysis, cultural heritage researchers can identify and prioritize preservation activities. Examples are restoration work done in Notre Dame Cathedral and analysis done on

the Terracotta Army (Ghaith, 2024). Moreover, by making use of AI tools in cultural preservation, future preservational needs can be identified and be planned for. Cultural heritage researchers have been using AI to aid cultural heritage preservation efforts. With AI applications, it has helped with identifying cultural heritage needs.

AI and machine learning systems have been used by cultural heritage researchers to decipher old scripts from Mayan ruins. With the use of deep learning technology, researchers can determine structural and grammatical rules for the system to transcribe and analyze the script (Ghaith, 2024). This brings light to previously inaccessible historical accounts for both academia and the general public. Furthermore, there have been developments by scholars to create digital and interactive interfaces with museums and cultural heritage sites. While it can facilitate access, it is important that researchers take into account data and privacy rights when it comes to intellectual property.

With the introduction of these AI systems in the humanities, it has created collaborative research opportunities among various social sciences fields. It is able to connect individuals from various fields of social sciences together. Some of these disciplines include linguistics, history, and computer science (Srivastava et al., 2025). With that being said, these advanced systems also come at a large price, and not all higher education institutions can afford them. This can cause issues with accessibility among institutions and create a potential for an AI-based digital divide. These challenges need to be considered so that there is proper oversight and equality across research universities.

AI sentiment and fact-checking tools are implemented within different fields of study across the humanities, from mental health research to research analyzing disinformation in the context of disease and cancer (Xu et al., 2024). Results from these AI models often correlate with data and events that are not AI-generated. It has been shown that AI sentiment analysis has been able to mimic the rises and falls of stock prices and consumer sentiment (Xu et al., 2024). While these sentiment and fact-checking AI tools are helpful, it is important to address issues to guarantee their validity in research and to avoid AI bias.

AI-powered and data-driven applications are widely used within digital humanities education, allowing students to receive personalized learning experiences. An area of study that heavily utilizes such systems is medical humanities. These technologies and educational methods show a high educational impact on a variety of specialized topics and across a variety of different students with different learning needs (Cahayani et al., 2024). It is important for stakeholders in education to address this digital divide because some teachers may lack the basic skills to operate these AI systems.

AI-implemented methods have had an impact on traditional interview practices. The integration of AI-led techniques, in combination with traditional approaches to interviewing in the humanities and social sciences, has helped researchers examine issues on a broader, more interconnected scale, enabling them to make generalizations to an array of groups and nations (Rezaev, Tregubova, 2018). This ability to increase the range of qualitative

investigation in the humanities helps to bring depth to the research conducted. However, care should be taken not to ignore the subtleties and specific contexts that could be revealed by a qualitative analysis (Rezaev, Tregubova, 2018).

AI has aided social sciences research in utilizing unstructured data. By using text and data analysis, it is helping to unveil new insights in fields such as behavioral research, public opinions, and trends in culture (Rolnik, 2024). It helps researchers to improve their analysis by making these large datasets smaller. AI has shown itself to be beneficial for social scientists. However, these benefits and the acceleration of studies using AI-based analyses must be examined in terms of ethical use and how to ensure that the tools do not distort human behavior.

AI is also transforming fields within the humanities by allowing for multilingual research through data processing of numerous languages using machine translations and AI-generative neural networks that create artworks from various data (Nanda et al., 2024). These AI applications are revolutionizing research methods within the humanities (Saddhono et al., 2024). They provide a different way for scholars to understand human behavior, which poses a number of theoretical implications within the field that will need to be examined by humanities researchers and scholars.

3.2. Theoretical Developments

The integration of AI into research has altered the construction of theoretical frameworks by discovering previously undetected patterns within qualitative data. The automation of the discovery of thematic clusters, latent variables, and patterns is making theory construction easier, providing chances to uncover latent associations and relationships between events or objects to reveal trends and relationships (Raj et al., 2023).

AI systems are capable of hypothesis generation and theory refinement, but these systems' scope of impact to construct complex constructs for sophisticated aspects of human behavior remains limited, and their ability to extrapolate theoretical concepts from data or analyze data in context also poses a question. By detecting empirical regularities, AI systems produce thematic clusters and latent variables, thereby serving as building blocks for data-informed theoretical frameworks that may lead to innovations or refinements in hypotheses (Christou, 2023). However, the overwhelming amount of computational findings may challenge researchers to find theoretical patterns of significance. As such tools for theory construction begin to garner popularity, it is imperative to question their validity and translatability, in diverse cultural or disciplinary circumstances, through reflective theorizing.

AI algorithms can be used for coding, which reduces the dependence on one researcher's ability to draw meaning out of textual data (Poldrack et al., 2023). The potential to produce standardized and replicable coding schemes has created a shift toward increased objectivity in the theoretical domain, but at the risk of inadvertently introducing biases into the theoretical development of phenomena of interest. As such, researchers must be cautious when applying algorithms that include data and programming biases that could affect or misrepresent

interpretation. For a theoretical framework to be built using data analytics, critical theorizing regarding the use of algorithms must be addressed. Theory-building in the social sciences requires situational context and the acknowledgement of variability, which may fall into the blind spots of algorithms.

AI-based visualization tools such as network graphs and word clouds show great potential to bring theoretical relationships in large datasets into more interpretable formations. Through visualization, AI systems open channels to discover new theories (Kovalerchuk et al., 2022). Theories, even those based on large amounts of quantitative data, are not inherently interpretive, and it is difficult to translate results into theoretical insights. Visualization methods are changing data-driven theories to become more interpretable and accessible across disciplines, which may oversimplify underlying theoretical nuances in order to clarify and convey relationships.

Computational methods and data-intensive approaches have changed traditional theory-driven methods through pattern discovery and uncovering correlations that might not be evident within pre-existing theoretical frameworks (Bircan, Salah, 2022). While these methods have brought new advancements, they also present a challenge to epistemology and have been labeled anti-theoretical since they prioritize uncovering patterns over causal interpretation. The inability to incorporate nuance, human agency, and context poses serious concerns regarding the development and relevance of theories. Critical theorizing in this respect requires reflective practices in how computational methods reframe or miss significant theoretical advances.

AI systems show promise in unveiling phenomena that are under-investigated or invisible within qualitative or small-sample-size research by uncovering patterns previously unseen (Qin, Zhang, 2025). Critical theorists must therefore reflect on the implications of how such pattern discoveries can extend or modify existing theories, as well as the necessity to reflect on theoretical concepts and generalizations. By being aware of how algorithms shape theoretical contributions, AI can reveal novel explanations, but theorists must be cautious to avoid overlooking important contextual or individual dynamics to ensure the proper adaptation and interpretation of new theoretical developments in order to capture the complexities of human behavior.

The ability of large language models (LLMs) to respond to questions with a human-like response allows for the generation of counterfactuals in order to explore theoretical possibilities, such as alternative outcomes to historical events or alternative responses among different demographics (Grossmann et al., 2023; Bail, 2024). By critically theorizing about how counterfactual analysis has the potential to shape theory development and the potential to ask "what-if" questions, the impact of AI systems can be maximized to inform theories for all dimensions of the social sciences (Xi et al., 2025). However, critical theorizing must reflect on how LLMs generate these models and if these models simply reconfigure biases embedded in the datasets used to train these AI tools. Theorizing must also be reflexive, reflecting on the

potential role of AI-generated content as evidence and whether theoretical concepts and analyses constructed using such evidence can lead to new and reliable theoretical generalizations.

LLMs show potential to uncover previously hidden sociocultural biases in theory generation, to theorize simulations that include more complex and complete situations, and to generate simulations more efficiently and at higher scale than with previous technologies, which can be used to examine diverse theoretical questions across the social sciences, expanding the horizons for research in areas such as small-group dynamics, public policy responses, or political attitudes (Yang et al., 2025). LLMs can simulate broad theoretical constructs across social groups, which are frequently limited in qualitative or small-sample studies. This shows the vast potential for LLMs to influence theoretical research. However, researchers must acknowledge the potential limitations and biases that AI tools can introduce, and consider how such limitations affect the interpretation of the data and the construction of theoretical generalizations, which is not addressed by the current research. Critical theorizing may reduce and control for the biases of the LLMs by critically examining the models they produce and using human interpretation to reflect the subtleties of complex human experiences in these simulations. Researchers must also reflect on the transparency and validity of LLMs, and consider ways to promote transparency and validity by considering the ethical challenges of LLMs in theory construction.

The development of data-based and theory-informed research methods has been further enhanced by the development of automated computer-based tools to evaluate large-scale cultural records, such as literary, historical, or archival records. Computer-based techniques are being applied to address research issues that are traditionally conceptualized, which leads to new connections between themes, interpretive frameworks, and aesthetic styles. The utilization of such large databases has also facilitated more comprehensive and empirically grounded reflections and debates about theories, and it can lead to new theoretical explanations within the social sciences and the humanities (Islam et al., 2024; McGillivray et al., 2022). A simple example of these methods being used in theory construction can be demonstrated with the patterns and trends detected in large datasets, which provide insights that shed new light on long-lasting interpretive issues. The availability of algorithms for theme or pattern extraction has provided unprecedented access to textual meaning, thereby adding an extra dimension to qualitative analysis (Gefen et al., 2020). While algorithm extraction of data enhances the accessibility of meaningful patterns, care must be taken to ensure that there is not an over-reliance on quantitative and empirical trends or patterns.

Theoretical concepts within the humanities are being constructed and utilized through pattern discovery of motifs, themes, and shifts in public debate within such cultural records as newspapers, literary works, and digitized artwork (Pawlowski, 2012; Smits, Wevers, 2023; Wijfjes, 2017).The use of computational tools provides the ability to investigate new or overlooked topics that would be impossible using manual data-mining techniques and to

examine these topics in tandem across multiple disciplines or genres. This calls for a model that reflects and addresses both digital humanities and social science domains, as both rely on human and computer interaction for successful theory-building. For these advancements in theory-building in the humanities to continue, critical theorizing on all levels is imperative. Theorizing must be both reflexive and critical of human interpretation in order to capture patterns that are of importance, so that theoretical research will not be overly focused or influenced by computer-extracted themes that overlook nuanced human experiences or interactions.

The ability of humans to produce explanations has led to the development of the subfield of AI called explainable AI (XAI) to explore meaningful explanations that go beyond mechanistic concepts (Xu et al., 2019). The development of XAI has been informed by philosophy and social psychology, which reflect on how to measure the qualities of a good explanation (Balog, Radlinski, 2020; Miller, 2019). Since XAI is human-centered, it is important to critical and reflective theorizing. Critical theorizing concerning theories of human understanding calls for reflexivity because human explanations are not isolated to algorithms, but are shaped by cognitive, behavioral, social, and cultural considerations. These elements, in combination with the impact of such forces on the production of algorithms and theories, impact the appropriateness and validity of new theories and how they are understood by human subjects.

The development of XAI has integrated human-centered models of explanation with algorithms, which have advanced and improved the performance and trustworthiness of algorithms (Miller, 2019). As a result, theorists must engage in reflective practice to consider the strengths and limitations of theoretical models, and they must evaluate and utilize AI tools critically and cautiously (Dwivedi et al., 2023). The advancements of XAI show the reciprocal relationship between AI and human understanding of theoretical explanation, in that AI methods can be enhanced using social science principles, and the use of AI methods and tools requires the reformulation of human theory explanation.

The use of AI in science has also led to the science of science, a field that leverages algorithms, machine learning, and AI to explore interconnected science networks with a view toward theoretical development on a meta level (Fortunato et al., 2018; Hou et al., 2025). Using large-scale and diverse datasets to conduct meta-analyses to understand, explain, and predict processes such as knowledge creation, diffusion, innovation, and impact calls for reflective theorizing on AI itself in order to prevent AI systems from replicating and reinforcing inequities, thereby reducing the system's validity and reliability. Data mining and mapping of millions of citations have discovered scientific macro-level trends of shifts in paradigms or the rate of diffusion of scientific knowledge (Han et al., 2022). Therefore, more complex and sophisticated theories must be conceptualized to account for these emergent trends, though theories must also consider the complexity and uniqueness of scientific behavior across disciplines.

Platforms such as SciScinet utilize machine learning techniques, computational algorithms, and visualization methods to provide comprehensive data-intensive methods to examine the co-evolution of science across the world and to identify and conceptualize new theories (Lin et al., 2023; Hou et al., 2025). The theoretical constructs conceptualized in the SciScinet project utilize the identification of quantitative shifts in the composition of scholarly communications and the integration of qualitative shifts in underlying knowledge as a foundation for assessing and forecasting paradigms. By incorporating computerization and automation to analyze knowledge, its creation, and scientific practices, scientists have contributed to advances in the theoretical development of science. However, the co-evolution of human behavior and AI in this realm must be further investigated and considered for the evolution of theory itself to be understood and predicted. Thus, theories of knowledge evolution should include the human-computer interaction as a necessary variable to understanding the dynamic influence of data and machines on how science changes.

The development of the content studies and content science fields has led to new theoretical concepts and frameworks regarding human activity and content activity in digital environments that go beyond and combine elements of established fields (Floridi, 2025; Lee et al., 2020). The field of content studies concerns itself with understanding the users, designers, developers, and consumers of content and how the environment impacts interaction, design, and navigation to provide meaningful immersive experiences for its users. Such a novel construction demands theories that account for the role of computational content and the dynamics of user-content interplay, and in doing so, theorists are addressing the lack of research and literature regarding user activity. By theorizing the user, content, and interactional phenomena and activities as well as how they co-evolve, research can provide a theoretical conceptualization for how content works in computational or non-computational spaces and how the user interacts with content to obtain meaningful information or immersion. However, critical theorizing remains absent in how theorizing and reflecting on the user-content dynamic or the co-evolution of user and content informs how content studies as a discipline has the responsibility to consider how the human-computer co-evolution also necessitates new ethical theories.

Critical theorizing in this endeavor remains untouched, even in a discipline that investigates human activity in such interactive mediums. In anthropological research, the examination of expert systems shows that theories are not only constructed but are embedded in the construction and utilization of AI systems (Bell, 2021). Anthropologists who have analyzed expert systems have emphasized the choices that programmers make while coding, and the limitations that such coding requires in order to make computer programs functional (Forsythe, 1993). Critical and reflective theorizing is necessary because while programmers may not be theorists by training or nature, their interpretation and categorization of human activity in theoretical fields as machine-readable is itself an act of theorizing and has theoretical implications. Researchers must be aware and cautious that while building computational tools,

they are contributing to how human knowledge is translated and utilized in systems that may or may not be able to capture human agency in an ethically just way.

AI holds promise in supporting critical theorizing by revealing unseen or unexplored areas of thought. Reflective theorizing and self-critique must, however, remain at the forefront of the practice and be utilized with AI, as the potential of AI as a method for aiding in the development of theory is both immense and has yet to be fully realized. The successful deployment of AI requires a comprehensive understanding of its capabilities and limitations, particularly in distinguishing between theoretical promise and practical implementation.

4. Ethical Considerations and Challenges

As AI technologies are gaining traction and impacting research practices in the social sciences and humanities, it is crucial to discuss the ethical considerations of data privacy, bias, and accountability in AI in order to ensure its responsible and transparent governance in the digital transformation of scholarship.

4.1. Data Privacy and Protection

The use of AI in the social sciences and humanities has undeniably increased the efficiency of research, but it also raised valid concerns in the protection of privacy. Automated systems, such as transcription or image processing software, have made it possible to process large amounts of data more accurately and quickly than before. However, these systems also increase the possibility of revealing sensitive participant information. For example, storing and processing interviews, surveys, and images on computers or cloud platforms makes data more accessible to unauthorized users (D'aquin et al., 2018; Pansoni et al., 2023). In the case of digitized and AI-analyzed manuscripts or artifacts, contextual clues linked to certain people or communities can be exposed, highlighting the need for the implementation of proper data governance (Neudecker, 2022). Moreover, the fact that AI technology is mainly automated increases the risks of data leakage, as large amounts of data are aggregated.

The algorithmic text processing and the use of AI models for cultural heritage have shown the importance of ethical and "privacy by design" practices (Birhane, 2021). AI tools that can conduct automated coding, entity recognition, and data enrichment often reveal information extracted from data that may have been supposedly anonymized beforehand (Tang, Su, 2024; D'aquin et al., 2018). For this reason, appropriate privacy controls and data audits should be regularly conducted to maintain appropriate data protection. However, ethical "privacy by design" processes are far from the status quo as institutions have not yet had the capacity, competence, and organizational will to put the framework to practice. In this respect, regular ethical updates and the institutional support of educational resources about ethical methodologies in these systems should be adopted (Runcan et al., 2025).

Digitalization processes in AI technology for cultural heritage have allowed the possibility of 3D scanning, mural restoration, and the construction of virtual museums, opening several questions about data privacy (Gervasi et al., 2022). The collection and use of high-resolution images and metadata can be potentially misused by indirectly providing information such as the identity of the artists or the community membership beyond what is initially required for the research (Ye, 2022; Pansoni et al., 2023). These issues are especially problematic when such cultural heritage can be related to social sensitivities. For that reason, beyond the legal compliance of these data, informed consent from members, cultural groups, and communities should be obtained to protect against a multitude of risks that could affect them, their dignity, privacy, and rights. As a result, ethical data guidelines in these cases are necessary for the interests and purposes of research.

While text and data recognition and enrichment techniques have helped in the information retrieval from library and cultural archives, they can sometimes lead to an ethical concern when personal historical data are unintentionally extracted or revealed. For example, the automated process of handwritten letters and historical documents can reveal insights about individuals, families, or communities without obtaining consent or any kind of partnership with stakeholders (Neudecker, 2022; Gonzalez, Rodrigues, 2022). These cases need to be addressed by creating transparent and comprehensive data curation policies, including the data selection, the data process, and the data usage for the purposes of the research. These policies should outline strategies for including communities in decisions about the curation of their own data, maintaining ethical data privacy.

Generative AI models, such as ChatGPT, may compromise data privacy, data protection, and intellectual property. With the utilization of personal and sensitive data by systems, data privacy is becoming a bigger issue within the research and educational fields. For example, an AI-driven system offering educational content or writing analysis could use personal data in its training process, consequently decreasing users' privacy (Nam, Bai, 2023; Farooq et al., 2023). Because of that, access controls and anonymization policies, among others, should be adopted in research and education to minimize ethical concerns by protecting individual and organizational interests.

Data-informed, ethical, and sustainable AI governance in research should implement cross-organizational initiatives and have a responsible leader. Ongoing concerns about accidental exposure, algorithmic opacity, and different privacy perceptions among researchers, communities, and organizations call for the evolution of these governance policies (Runcan et al., 2025). Reporting transparency, the implementation of continuous training, and the development of organizational cultures for AI-aided data privacy are necessary to ensure more ethical and equitable uses of these technologies. Also, integrating community perspectives into the creation of research guidelines can increase policy uptake and accountability.

In conclusion, as AI is becoming more prevalent in the social sciences and the humanities, researchers, organizations, and stakeholders should make more efforts in the implementation of ethical policies regarding data privacy and protection.

4.2. Bias and Accountability

AI tools have brought in innovative approaches for social sciences and humanities research (Frontoni et al., 2024; Balburin et al., 2024). However, their potential to exacerbate bias, undermine accountability, and lack transparency remain prominent concerns that require deliberate mitigation. The replication and reinforcement of societal biases by AI-driven models remain a critical issue (Khan, 2023). These models are trained on historical data, which inevitably mirrors the structural inequalities that exist within societies (Ho et al., 2025). As a result, these models often perpetuate or even exacerbate the replication of biases such as race, gender, and class, leading to discrimination in research outcomes and social policies (Lindgren, Holmström, 2020; Tang, Su, 2024). Active counterbalancing, starting with bias mitigation through diversified and ethical training datasets, is crucial.

When AI outputs in research are biased, it impacts decision-making in policy and the way marginalized groups are depicted and described. For instance, sentiment analysis, based on biased data, can distort public opinion and cultural representations (Celi et al., 2022). This negatively impacts not only the integrity of research but also the quality of policy recommendations that emerge. To combat this, implementing robust auditing tools, as well as the investigation and implementation of fairness-aware algorithms to balance outcomes between groups, are deemed effective interventions to deal with bias in AI applications.

Minimizing bias necessitates increasing demographic diversity of researchers and decision-makers, given that heterogeneity has shown to promote equity in the distribution of outcomes (Messeri, Crockett, 2024). Such diversity encourages a wide range of perspectives on how AI systems can be developed and used in diverse contexts, ensuring that such systems do not solely perpetuate traditional norms. To build an inclusive research workforce, universities and research institutions must make recruitment, training, and collaboration strategies more inclusive to ensure that the diversity of the communities they serve is represented in AI research and decision-making. Additionally, the involvement of underrepresented groups in this process helps promote accountability and can ensure the creation and use of AI systems for socially and economically relevant purposes.

The insufficient AI training of social science and humanities researchers puts the validity and credibility of AI application at risk, due to the researchers' low technical competence in AI methods, making them adopt and reproduce AI outputs more uncritically (Lavidas et al., 2024; Chakravorti et al., 2024). For this, researchers need targeted training in AI literacy, which could be facilitated via focused workshops or interdisciplinary educational approaches, where these scholars are offered adequate critical techniques to evaluate and contextualize AI output, thus ensuring more reliable and socially responsive research findings.

The opacity of AI systems further hinders accountability, as the models often have internal functioning that is opaque and unintelligible for researchers and end-users to understand. Such opacity of algorithms in the AI domain, especially for machine learning methods, undermines trust in AI-driven results, as users cannot be aware of whether such results reflect methodological issues or reflect a lack of transparency (Tang, Su, 2024). Disclosure of all used algorithms and AI tools, datasets, and parameters is necessary so as to improve transparency and enables comprehensive evaluation and auditing of AI systems (Runcan et al., 2025).

Explainable AI (XAI) allows end-users to see the reasoning behind AI predictions, which is particularly helpful in explaining how AI systems arrive at certain conclusions. This bridges the gap between human and technological interpretation (Miller, 2019), therefore addressing AI opacity. The interpretability in XAI is ensured if explanations are produced for humans (Tull et al., 2024). This demands interdisciplinary cooperation and collaborative work between software and AI engineers and the community and stakeholders, to provide qualitative interpretations of XAI's results (Miller, 2019).

Interpretive misrepresentation occurs when analyzing humanities and cultural datasets and archives, which have been digitized by algorithms. In humanities research, context, meaning, and interpretation depend heavily on a human researcher's subjective and historical perception. Thus, there is a risk of losing sight of the subjectiveness of humanities and cultural studies as research becomes increasingly dependent on algorithms. Algorithmic-based humanities research can potentially overlook or misinterpret the meanings and subjective elements embedded in sources and archives, as AI focuses on patterns and trends in historical text that often lack cultural relevance. Therefore, the deployment of interdisciplinary ethics committees in AI research is required, which, alongside participatory governance and institutional codes, can oversee accountability and decision-making for AI use (Runcan et al., 2025).

The overuse of algorithmization could risk deprioritizing human interpretation, which is at the core of humanities research, which can diminish the values of context-dependent interpretations and evaluations of datasets (Leeuw, 2025). Research can become largely focused on computational efficiency without adequately exploring the contextual and qualitative nuances of the datasets, for example (Frontoni et al., 2024). Algorithmization has become the dominant form of data analysis because of the increased computing power (Frontoni et al., 2024), therefore undermining the human interpretative role, as research now is focused on quantitative analyses. This phenomenon is particularly detrimental for those disciplines or types of research that are reliant on non-algorithmic qualitative analyses.

The dominance of a mainstream cultural dataset risks neglecting marginalized groups, thus replicating power imbalances within the academic field. A lack of critical awareness of which datasets are used to train algorithmic models may promote the dominance of those with more mainstream representation and historically powerful or dominant voices (Lindgren, Holmström, 2020). For this reason, the integration of human expertise at all steps of algorithmic research is vital to ensure inclusivity in knowledge generation and model design (Frontoni et al., 2024).

Some of the studies in our sample debate whether to automate all the research phases, including data interpretation, or if AI algorithms are suitable for data processing, but interpretation needs to be done by the researcher or by an AI program adapted for the humanities or social sciences field (Chun, Elkins, 2023). Therefore, research needs to establish clear boundaries in which it is specified whether algorithms should aid or substitute interpretive work and how research can incorporate humans in all the research phases (Frontoni et al., 2024).

Although some researchers show concerns that algorithmic approaches will diminish the interpretative capacity of scholars or over-rely on mathematical analyses, other studies in our sample propose the possibility of hybrid interpretations to mediate this tendency in which the results are produced by the computer, and interpreted by the researcher or a computer program, allowing for the simultaneous application of interpretative analysis and algorithmic modeling in a way that does not interfere with one another.

Familiarity levels with AI differed between younger (doctoral candidates) and older (established academics) researchers. Doctoral candidates tend to embrace and display confidence with new AI technologies, while senior academic staff were less AI literate and skilled, and thus could not oversee and supervise doctoral research to the degree required. To improve transparency and interpretability, it is important to have more training workshops for researchers so they will have AI literacy across disciplines, as well as that universities have to develop long-term, continuing-capacity learning programs to cater for both senior academic staff, junior researchers, and doctoral students (Lavidas et al., 2024).

The rapid evolution of AI technologies necessitates the development of comprehensive educational frameworks that prepare future scholars and practitioners for the ethical and effective application of these tools (Messeri, Crockett, 2024). Furthermore, future research must proactively address the inherent biases within AI training data, considering how these biases can be leveraged not as flaws, but as crucial insights into historical interpretation and socio-political forces (Baunvig et al., 2025). This includes actively working towards the democratization of data and ensuring human accountability in AI decision-making processes to mitigate potential discrimination and uphold individual autonomy.

The ethics governance in our sample frames the role of researchers as "stewards of AI" by establishing AI ethics protocols, which are built on institutional codes and community narratives of appropriate use (Jeon et al., 2025). These ethics frameworks provide transparency of procedures and can encourage accountability among AI creators and researchers to minimize AI bias, which creates trust and confidence among the public, to further reduce miscommunication and fear caused by uncertainty (Jeon et al., 2025; Runcan et al., 2025).

AI governance has to consider institutional resources and external resources available and be adapted to specific needs, values, and contexts. Governance through multi-stakeholder involvement requires that institutions incorporate societal, cultural, and ethical values into all steps of the development and use of AI for social sciences and humanities research (Cai et al., 2019). Effective governance frameworks are essential to navigate these complexities, ensuring that AI development and deployment align with ethical principles and societal well-being.

5. Summary

This paper has explored the multifaceted integration of AI within the social sciences and humanities, highlighting both its transformative potential as a research tool and its complex implications for ethical considerations and societal impact. The study aimed to critically assess how artificial intelligence is changing research in the social sciences and humanities. It sought to evaluate opportunities, challenges, and innovations in methods, theories, and ethics brought by AI integration. The literature review and synthesis of case studies have accomplished the objective, showing the many ways AI is impacting the social sciences and humanities, in both research and ethical considerations.

AI technologies are transforming research methodology in many disciplines in the social sciences and humanities. Automated data processing has facilitated quantitative and qualitative data analysis in ways never before possible. Pattern recognition and natural language processing are examples of how AI tools can advance thematic analysis and make sense of vast amounts of complex and unstructured information.

Ethical considerations are vital when discussing the implementation of AI into research. AI technologies offer new possibilities for research methods and theory development; however, they have increased ethical concerns around privacy, data security, algorithmic bias, and accountability. The potential for AI to perpetuate or even amplify existing societal biases, particularly through the data used for training machine learning algorithms, represents a critical ethical challenge that demands rigorous attention (Leavy et al., 2020). This includes addressing biases inherent in AI algorithms and ensuring the responsible use of AI to protect privacy and promote fairness in research (Markowitz et al., 2024). This requires rigorous evaluation of AI systems to mitigate risks such as algorithmic opacity, which can limit accountability and contestability, and the potential for perpetuating inequalities through automated decision-making (Radavoi, 2025). Moreover, the social alienation generated by AI, stemming from job displacement and reduced human interaction, also warrants careful consideration within social research.

When situating the findings of this study in the larger context of existing literature, it appears that AI technologies are creating an impact by altering traditional social science and humanities methods. While there is limited available data on how AI-based methodologies may be implemented and how social science and humanities scholars can benefit from automated technologies, AI has greatly increased the possibilities of interdisciplinary work. AI offers more efficient data analysis and the capability of testing complex hypotheses, enabling new interdisciplinary and cross-disciplinary exploration and theory development.

Some limits to this study were considered throughout the research, as it is impossible to create a study that perfectly captures the complexity of the interaction of humans and AI in a research setting. Some considerations and limitations are summarized below:

- Some information and data may become outdated. The speed at which AI technologies are emerging makes any research project on this topic dynamic.
- The data from this study came exclusively from secondary sources. Though many different sources were consulted, there was no empirical data collection as a part of this study. This may be an important limitation.

Despite the possible limitations, this study has been important in assessing how AI technology is altering methods in the social sciences and humanities as well as ethical concerns in this realm. Some directions for future research on AI in the social sciences and humanities may be explored in the following section.

Future research is critical to provide more detailed understanding on the incorporation of AI technologies into social sciences and humanities scholarship. Three key areas for future research are outlined below:

- Empirical research on methods. This research would assess how AI has changed theoretical assumptions and how it is being applied to practical methods. It would be valuable for this study to analyze both quantitative and qualitative findings on various AI methodologies in order to assess how scholars are developing new strategies for implementing AI into both methodological and theoretical research.
- Future research on ethics of AI in social sciences and humanities research. This is an underrepresented area in the literature. To maintain accountability, ethical considerations need to be addressed to ensure that scholars are being responsible and respectful with both data and humans as AI-based methodologies are implemented.
- More complex empirical research of AI methodology in the social sciences and humanities will also be needed to address gaps in understanding of various types of methods and how they can work together. For example, how AI-based pattern recognition can be implemented into a study that is employing methods related to ethnographic inquiry.

Though the future of research and literature is speculative and unconfirmed, the directions above are some of the most helpful and worthwhile paths toward deeper comprehension of these technologies. Two recommendations emerge from the above section. These are:

- Further encourage discussion and research on ethical standards that need to be put in place to ensure responsible use of AI technologies.
- Continued discussions on how to incorporate and integrate various techniques from AI-based research into traditional social science and humanities-based methods.

AI significantly enhances the study of digital humanities by enabling advanced computational analyses of extensive cultural datasets, offering unprecedented insights into linguistic patterns, historical narratives, and artistic expressions. For instance, AI-powered tools can meticulously analyse vast textual corpora to identify thematic trends and stylistic nuances that might be imperceptible through traditional human-centric methods. This analytical

capability extends to the preservation of cultural heritage, where AI can aid in the digital reconstruction of damaged artifacts and the cataloging of extensive archives.

The fields of social sciences and humanities, traditionally centered on human experience, culture, and societal structures, are now poised to leverage AI's analytical capabilities to explore complex human and social problems, offering new avenues for research into subjective experiences and interactive activities (Huang, 2022). AI, a technology capable of reproducing human-like reasoning, learning, planning, and creativity through data processing and algorithmic responses, provides a powerful tool for empirical research, particularly for prediction and recognition tasks involving vast datasets (Rascão, 2024). This integration marks a transition towards a "digital empirical approach", where research models and analyses are increasingly data-driven, representing a significant evolution from traditional qualitative and quantitative methodologies.

References

1. Arulselvi, V. (2024). Exploring AI for Social Science Research: Key Tools and Applications. *Artificial Intelligence, 29*, pp. 58-75.
2. Baiburin, A., Berezkin, Y., Gromov, A., Kovalenko, K., Sokolov, E., Kovalyova, N., Moskvitina A. et al. (2024). Artificial Intelligence in the Social Sciences and Humanities. In: *Forum for Anthropology and Culture, no. 20*, pp. 11-60. https://doi.org/10.31250/1815-8870-2024-20-20-11-60
3. Bail, C.A. (2024). Can Generative AI improve social science? *PNAS, 121(21),* e2314021121. https://doi.org/10.1073/pnas.2314021121
4. Balog, K., Radlinski, F. (2020, July). *Measuring recommendation explanation quality: The conflicting goals of explanations.* Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, pp. 329-338, https://doi.org/10.1145/3397271.3401032
5. Baunvig, K.F., Tafdrup, J., Vad, K., Rasmussen, K.S.G., Väina, M. (2025). From Bias to Insight : Computational Challenges and Opportunities in the Humanities. *Research Portal Denmark, 33*. https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=au&id=au-ceb30ae6-78b2-4acf-9fd4-890584aaeed2&ti=From%20Bias%20to%20Insight%20%3A%20Computational%20Challenges%20and%20Opportunities%20in%20the%20Humanities
6. Bell, G. (2021). Talking to AI: An anthropological encounter with artificial intelligence. *The SAGE handbook of cultural anthropology, 1*, pp. 442-458.
7. Bircan, T., Salah, A.A.A. (2022). A bibliometric analysis of the use of artificial intelligence technologies for social sciences. *Mathematics, 10(23),* 4398. https://doi.org/10.3390/math10234398

8. Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. *Patterns, 2(2).* https://doi.org/10.1016/j.patter.2021.100205
9. Cahayani, W.A., Rahayu, I.D., Purwantiningrum, D.A. (2024). A digital humanities exploration of AI-driven learning technologies in medical education. *Advances in Economics, Business and Management Research, 294,* pp. 414-420. https://doi.org/10.2991/978-94-6463-525-6_48
10. Cai, Y., Ramis Ferrer, B., Martinez Lastra, J.L. (2019). Transnational innovation ecosystems: Towards a transdisciplinary approach of integrating social sciences and artificial intelligence. *Sustainability, 11(17),* 4633. https://doi.org/10.3390/su11174633
11. Celi, L.A., Cellini, J., Charpignon, M.L., Dee, E.C., Dernoncourt, F., Eber, R., ..., Yao, S. (2022). Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLOS digital health, 1(3),* e0000022
12. Chakravorti, T., Wang, X., Venkit, P.N., Koneru, S., Munger, K., Rajtmajer, S. (2024). *Social scientists on the role of AI in research*. Association for the Advancement of Artificial Intelligence. https://www.arxiv.org/pdf/2506.11255
13. Chapinal-Heras, D., Díaz-Sánchez, C. (2023). A review of AI applications in Human Sciences research. *Digital Applications in Archaeology and Cultural Heritage, 30*, e00288. https://doi.org/10.1016/j.daach.2023.e00288
14. Christou, P.A. (2023). The use of artificial intelligence (AI) in qualitative research for theory development. *The Qualitative Report, 28(9),* pp. 2739-2755. https://philpapers.org/archive/CHRTUO-3.pdf.
15. Chun, J., Elkins, K. (2023). The crisis of artificial intelligence: A new digital humanities curriculum for human-centred AI. *International Journal of Humanities and Arts Computing, 17(2),* 147-167. https://doi.org/10.3366/ijhac.2023.0310
16. D'aquin, M., Troullinou, P., O'Connor, N., Cullen, A., Faller, G., Holden, L. (2018). Towards an "Ethics by Design" methodology for AI research projects. *AAAI/ACM Conference on AI, Ethics, and Society,* 54-59. https://hal.science/hal-03650299/document.
17. Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., ..., Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM computing surveys, 55(9),* 1-33. https://doi.org/10.1145/3561048
18. Farooq, M., Buzdar, H.Q., Muhammad, S. (2023). Al-Enhanced social sciences: A systematic literature review and bibliographic analysis of Web of Science published research papers. *Pakistan Journal of Society, Education and Language (PJSEL), 10(1),* 250-267.
19. Floridi, L. (2025). Content studies: A new academic discipline for analysing, evaluating, and designing content in a digital and AI-driven age. *Philosophy & Technology, 38*, 1-17. https://doi.org/10.1007/s13347-025-00877-6.
20. Forsythe, D.E. (1993). Engineering knowledge: The construction of knowledge in artificial intelligence. *Social Studies of Science, 23(3),* pp. 445-477.

https://doi.org/10.1177/0306312793023003002

21. Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., ..., Barabási, A.L. (2018). Science of science. *Science, 359(6379),* eaao0185. https://doi.org/10.1126/science.aao0185.
22. Frontoni, E., Paolanti, M., Migliorelli, L., Pietrini, R., Asimakopoulos, S. (2024). Artificial intelligence: The new frontier in digital humanities. *Frontiers in Computer Science, 6*, 1-17. https://doi.org/10.3389/fcomp.2024.1529826
23. Gefen, A., Saint-Raymond, L., Venturini, T. (2020). AI for Digital Humanities and Computational Social Sciences. *Reflections on AI for Humanity.* Springer, 1-13.
24. Gervasi, O., Perri, D., Simonetti, M., Tasso, S. (2022, July). *Strategies for the digitalization of cultural heritage.* International Conference on Computational Science and Its Applications Cham: Springer International Publishing, pp. 486-502. https://doi.org/10.1007/978-3-031-10592-0_35
25. Ghaith, K. (2024). AI integration in cultural heritage conservation – Ethical considerations and the human imperative. *International Journal of Emerging and Disruptive Innovation in Education,* 1-8. https://digitalcommons.lindenwood.edu/cgi/viewcontent.cgi?article=1022&context=ijedie
26. Gonzalez, M.E., Rodrigues, M.V. (2022). Digital Humanities: Ethical implications and interdisciplinary challenges. *Humanities Bulletin, 5(1),* 111-125. http://journals.lapub.co.uk/index.php/HB/article/download/2357/1622
27. Grossmann, I., Feinberg, M., Parker, D.C., Christakis, N.A., Tetlock, P.E., Cunningham, W.A. (2023). AI and the transformation of social science research. *Science, 380(6650),* 1108-1109. https://humannaturelab.net/sites/default/files/2024-04/science.adi1778.pdf
28. Han, J., Pei, J., Tong, H. (2022). *Data mining: concepts and techniques*. Morgan Kaufmann.
29. Ho, J.Q., Hartanto, A., Koh, A., Majeed, N.M. (2025). Gender biases within Artificial Intelligence and ChatGPT: Evidence, sources of biases and solutions. *Computers in Human Behavior: Artificial Humans, 100145*. https://doi.org/10.1016/j.chbah.2025.100145
30. Ho, M.T., Vu, T.H.T. (2025). Cultural analytics amid the rise of generative AI: critical insights for human-AI cocreative cultural studies. *AI & SOCIETY,* 1-4. https://doi.org/10.1007/s00146-025-02489-2
31. Hou, J., Zheng, B., Li, H., Li, W. (2025). Evolution and impact of the science of science: from theoretical analysis to digital-Al driven research. *Humanities & Social Sciences Communications, 12(316),* 1-9. https://doi.org/10.1057/s41599-025-04617-1
32. Huang, Z. (2022). Introducing Neuro-Symbolic Artificial Intelligence to Humanities and Social Sciences: Why Is It Possible and What Can Be Done? *TEM Journal, 1863.* https://doi.org/10.18421/tem114-54
33. Hutson, J. (2024). Rethinking plagiarism in the era of generative AI. *Journal of Intelligent Communication, 4(1).* https://digitalcommons.lindenwood.edu/faculty-research-papers/619?utm_source=digitalcommons.lindenwood.edu%2Ffaculty-research-

papers%2F619&utm_medium=PDF&utm_campaign=PDFCoverPages
34. Islam, M.N., Hu, G., Ashiq, M., Ahmad, S. (2024). *Exploring the landscape of big data applications in librarianship: A bibliometric analysis of research trends and patterns*. Library Hi Tech. https://doi.org/10.1108/LHT-05-2023-0193
35. Jeon, J., Kim, L., Park, J. (2025). The ethics of generative AI in social science research: A qualitative approach for institutionally grounded AI research ethics. *Technology in Society, 81,* 1-3. https://doi.org/10.1016/j.techsoc.2025.102836
36. Khan, S. (2023). The Ethical Imperative: Addressing bias and discrimination in AI-driven education. *Social Sciences Spectrum, 2(1),* pp. 89-96.
37. Köntges, T. (2020). Measuring philosophy in the first thousand years of Greek literature. *Digital Classics Online*, 1-23. https://doi.org/10.11588/dco.2020.2.73197
38. Kovalerchuk, B., Nazemi, K., Andonie, R., Datia, N., Banissi, E. (Eds.) (2022). *Integrating artificial intelligence and visualization for visual knowledge discovery, Vol. 1014*. Springer Nature.
39. Lavidas, K., Voulgari, I., Papadakis, S., Athanassopoulos, S., Anastasiou, A., Filippidi, A., Komis, V., Karacapilidis, N. (2024). Determinants of Humanities and Social Sciences Students' Intentions to Use Artificial Intelligence Applications for Academic Purposes. *Information, 15(6),* 314. https://doi.org/10.3390/info15060314
40. Leavy, S., O'Sullivan, B., Σιαπέρα, E. (2020). *Data, Power and Bias in Artificial Intelligence.* arXiv (Cornell University). https://doi.org/10.48550/arxiv.2008.07341
41. Lee, L.W., Dabirian, A., McCarthy, I.P., Kietzmann, J. (2020). Making sense of text: artificial intelligence-enabled content analysis. *European Journal of Marketing, 54(3),* pp. 615-644. https://doi.org/10.1108/EJM-02-2019-0219
42. Leeuw, F.L. (2025). The Algorithmization of Policy and Society. *Artificial Intelligence and Evaluation, 242.* Taylor & Francis.
43. Lin, Z., Yin, Y., Liu, L., Wang, D. (2023). SciSciNet: A large-scale open data lake for the science of science research. *Scientific Data, 10(1),* 315. https://doi.org/10.1038/s41597-023-02198-9
44. Lindgren, S., Holmström, J. (2020). A social science perspective on artificial intelligence: Building blocks for a research agenda. *Journal of Digital Social Research, 2(3),* 1-15. https://publicera.kb.se/jdsr/article/download/25135/20308
45. Mahoney, J. (2021). *The logic of social science*. Princeton University Press.
46. Markowitz, D.M., Boyd, R.L., Blackburn, K. (2024). From silicon to solutions: AI's impending impact on research and discovery. *Frontiers in Social Psychology, 2.* https://doi.org/10.3389/frsps.2024.1392128
47. McGillivray, B., Marongiu, P., Pedrazzini, N., Ribary, M., Wigdorowitz, M., Zordan, E. (2022). Deep impact: A study on the impact of data papers and datasets in the humanities and social sciences. *Publications, 10(4),* 39. https://doi.org/10.3390/publications10040039.
48. Messeri, L., Crockett, M.J. (2024). Artificial intelligence and illusions of understanding in

scientific research. *Nature, 627*, pp. 49-58. https://doi.org/10.1038/s41586-024-07146-0

49. Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1-38. https://doi.org/10.1016/j.artint.2018.07.007
50. Mojadeddi, Z., Rosenberg, J. (2024). Automated transcription of interviews in qualitative research using artificial intelligence: a simple guide. *J. Surg. Res. Prac.*, *5*, 1-6. https://doi.org/10.46889/JSRP.2024/5204
51. Nam, B.H., Bai, Q. (2023). ChatGPT and its ethical implications for STEM research and higher education: A media discourse analysis. *International Journal of STEM Education, 10(66),* 1-24. https://doi.org/10.1186/s40594-023-00452-5
52. Nanda, S.K., Panda, P., Satpathy, S. (2024). Exploring the synergy: Impact of artificial intelligence on humanities. *YMER, 23(02),* pp. 220-226. https://ymerdigital.com/uploads/YMER230240.pdf
53. Neudecker, C. (2022). Cultural Heritage as Data: Digital Curation and Artificial Intelligence in Libraries. *CEUR Workshop Proceedings, Vol. 3234*, 1-6. http://ceur-ws.org/Vol-3234/paper2.pdf
54. Ngalande, S. (2011). A logical analysis of selected texts in Nyanja. *African Study Monographs, 32(3),* pp. 91-109.
55. Ngulube, P., Mathipa, E.R., Gumbo, M.T. (2015). Theoretical and conceptual frameworks in the social and management sciences. In: *Mathipa ER and Gumbo MT (eds) Addressing Research Challenges: Making Headway for Developing Researchers* (pp. 43-66). Noordwyk, South Africa: Mosala-MASEDI.
56. Pansoni, S., Tiribelli, S., Paolanti, M., Di Stefano, F., Frontoni, E., Malinverni, E.S., Giovanola, B. (2023). Artificial intelligence and cultural heritage: Design and assessment of an ethical framework. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48(M-2),* pp. 1149-1155. https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1149-2023
57. Pawlowski, T. (2012). *Concept formation in the humanities and the social sciences, Vol. 144.* Springer Science & Business Media.
58. Peregrin, J., Svoboda, V. (2017). *Reflective equilibrium and the principles of logical analysis: Understanding the laws of logic.* Taylor & Francis.
59. Poldrack, R.A., Lu, T., Beguš, G. (2023). *AI-assisted coding: Experiments with GPT-4.* arXiv preprint arXiv:2304.13187. https://doi.org/10.48550/arXiv.2304.13187
60. Qin, Q., Zhang, S. (2025). Visualizing the knowledge mapping of artificial intelligence in education: A systematic review. *Education and Information Technologies, 30(1),* pp. 449-483. https://doi.org/10.1007/s10639-024-13076-1
61. Radavoi, C.N. (2025). Preserving human relevance, as a new social responsibility of business in the AI age. *Social Responsibility Journal*. https://doi.org/10.1108/srj-01-2025-0011

62. Raj, S., Jamthe, S., Viswanath, Y., Lokiah, S. (2023). *Introducing Construct Theory as a Standard Methodology for Inclusive AI Models.* arXiv preprint arXiv:2304.09867. https://doi.org/10.48550/arXiv.2304.09867
63. Rascão, J. (2024). Does Digital Truth correspond to Human Truth in the Digital Society (from Theory to Practice)? *American Journal of Humanities and Social Sciences Research (AJHSSR), Vol. 8, Iss. 9*, pp. 217-258.
64. Rezaev, A.V., Tregubova, N.D. (2018). Sociology as a science: New challenges of the 21st century. *Monitoring of Public Opinion: Economic and Social Changes, 5,* pp. 91-108. https://doi.org/10.14515/monitoring.2018.5.10
65. Rolnik, Z. (2024). The impact of artificial intelligence on academic research. *Universal Library of Innovative Research and Studies, 1(1),* 9-11. https://doi.org/10.70315/uloap.ulirs.2024.0101002
66. Runcan, R., Hațegan, V., Toderici, O., Croitoru, G., Gavrila-Ardelean, M., Cuc, L. D., Rad, D., Costin, A., Dughi, T. (2025). Ethical AI in social sciences research: Are we gatekeepers or revolutionaries? *Societies, 15(3),* pp. 1-17. https://doi.org/10.3390/soc15030062
67. Saddhono, K., Rohmadi, M., Lestari, T.A., Simanungkalit, K.E., Sukmono, I.K. (2024, November). *The Role of AI in Facilitating Multilingual Literacy and Cross-Cultural Understanding.* 2024 International Conference on IoT, Communication and Automation Technology (ICICAT), pp. 366-371.
68. Smits, T., Wevers, M. (2023). A multimodal turn in Digital Humanities. Using contrastive machine learning models to explore, enrich, and analyze digital visual historical collections. *Digital Scholarship in the Humanities, 38(3),* pp. 1267-1280. https://doi.org/10.1093/llc/fqad008
69. Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research, 104,* pp. 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039
70. Srivastava, P., Choudhary, R.R., Tekwani, K., Srivastava, A. (2025). Implementation of AI in humanities and social sciences. In: Taylor & Francis Group (Ed.), *Recent advances in sciences, engineering, information technology & management (1st ed.)* (p. 7). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003598152-62/implementation-ai-humanities-social-sciences-prerna-srivastava-ritu-raj-choudhary-kritika-tekwani-ankur-srivastava
71. Tang, L., Su, Y.S. (2024). Ethical implications and principles of using artificial intelligence models in the classroom: A systematic literature review. *International Journal of Interactive Multimedia and Artificial Intelligence, 8(5),* pp. 25-36. https://doi.org/10.9781/ijimai.2024.02.010.
72. Tull, S., Lorenz, R., Clark, S., Khan, I., Coecke, B. (2024). *Towards compositional interpretability for xai.* arXiv preprint arXiv:2406.17583. https://doi.org/10.48550/arXiv.2406.17583.

73. Unke, O.T., Chmiela, S., Sauceda, H.E., Gastegger, M., Poltavsky, I., Schutt, K.T., ..., Muller, K.R. (2021). Machine learning force fields. *Chemical Reviews*, *121*(16), pp. 10142-10186. https://dx.doi.org/10.1021/acs.chemrev.0c01111?ref=pdf
74. Wijfjes, H. (2017). Digital humanities and media history: A challenge for historical newspaper research. *Tijdschrift voor Mediageschiedenis, 20(1),* 4. https://doi.org/10.18146/2213-7653.2017.277
75. Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ..., Gui, T. (2025). The rise and potential of large language model based agents: A survey. *Science China Information Sciences, 68(2),* 121101. https://doi.org/10.1007/s11432-024-4222-0
76. Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., Zhu, J. (2019, September). *Explainable AI: A brief survey on history, research areas, approaches and challenges.* CCF international conference on natural language processing and Chinese computing. Cham: Springer International Publishing, pp. 563-574, https://doi.org/10.1007/978-3-030-32236-6_51
77. Xu, R., Sun, Y., Ren, M., Guo, S., Pan, R., Lin, H., Sun, L., Han, X. (2024). *AI for social science and social science of AI: A survey.* Elsevier. https://arxiv.org/pdf/2401.11839
78. Yang, S., Krause, N.M., Bao, L., Calice, M.N., Newman, T.P., Scheufele, D.A., ..., Brossard, D. (2025). In AI we trust: The interplay of media use, political ideology, and trust in shaping emerging AI attitudes. *Journalism & Mass Communication Quarterly, 102(2),* pp. 382-406. https://doi.org/10.1177/10776990231190868
79. Ye, J. (2022). The application of artificial intelligence technologies in digital humanities: Applying to Dunhuang culture inheritance, development, and innovation. *Journal of Computer Science and Technology Studies, 4(2.5).* https://doi.org/10.32996/jcsts. 2022.4.2.5
80. Yu, S., Carroll, F., Bentley, B.L. (2024). Insights into privacy protection research in AI. *IEEE Access, 12*, 41704-41726.10.1109/ACCESS.2024.337812