IJIS: Vol.2, Issue 1, January 2026 Page: 75-88 ISSN 3049-3277

International Journal of Integrative Studies (IJIS)
Journal homepage:www.ijis.co.in

Artificial Intelligence in Research Methodology: Methodological Advances, Opportunities, Risks, and Responsible AI Integration Framework

Conference Article

Ch. Mohammad Akram
Research Scholar, Department of Mechanical Engineering,
Acharya Nagarjuna University, Guntur, Andhra Pradesh – 522510, India.
akram.anucet@gmail.com

Abstract

This article discusses the integration of artificial intelligence (AI) as part of the research methodology and studies its methodological implications, potential opportunities, advancements, and threats. The study quantifies how useful AI can be in automating complex, broad, and long chain syllogistic operations and computations, as well as accelerating the predictions and projections in data-intensive domains. The study, which uses a mixed-method design that combines empirical assessments of the literature and practitioner-based case studies of four varied applications of AI in research, namely text analytics, sentiment analysis, climate modelling, and health-care research, finds that the AI-based workflows evaluated by the study are able to achieve significant gains in efficiencies and accuracies with remarkably similar patterns. The results also pose potential concerns, such as algorithmic bias, limited explainability and reproducibility, and other ethical and data-privacy risks. Our research outlines a framework and governance mechanisms to enable the responsible integration of AI to overcome challenges. The research cycle benefits from these mechanisms through the injection of bias mitigation and transparency. This article summarizes some of the key guidelines on the ethical use of AI that ensure transparency and rigor.

Keywords: Artificial Intelligence (AI); Research Methodology; Responsible AI; Explainable Artificial Intelligence (XAI); Algorithmic Bias; Research Governance; Mixed-Method Research; Aspect-Based Sentiment Analysis (ABSA).

1. Introduction

Modern research methods typically include artificial intelligence, also known as AI. Technologies of artificial intelligence (AI) are being deployed increasingly to collect, scrutinize, interpret, and synthesize information in a data-rich world *[1–3]*. Recent developments in AI, machine learning, deep learning, natural language processing (NLP), and more make it possible for researchers to automate syndrome analytical activities, handle large volumes of unstructured data, and make more accurate predictions in information systems, healthcare, climate science, and social sciences *[4–7]*. As researchers embrace artificial intelligence (AI), they must reflect on methodological and governance concerns. However, a lack of consideration for the latter is noticeable. Many papers discuss the augmentation of performance gains (e.g., efficiency gains, accuracy enhancements, etc.) at an application level and do not, and critically, reflect on the effects of AI on fundamental research standards (e.g., validity, transparency, reproducibility, accountability, etc.) *[8–12]*. Such is the case, for instance, of the application of generative AI and large language models (LLMs) for academic research and knowledge development *[4,13,14]*. AI-enabled investigative methods are being cautioned against by rising research that suggests high-risk usage. Algorithmic bias has now become a documented reality, through unrepresentative, incomplete, or historically biased training *[15–17]*. Due to their complexity, ML methods do not justify their results or interpretations of their findings *[18–21]*. In addition, the presence of ethical issues, including those related to privacy, accountability, and governance, further complicates research practice in AI adoption *[22–26]*. This paper will critically analyze the Artificial Intelligence methods to counter these challenges in resolving the issues. The study, based on the mixed-method evidence from various empirical case studies and practitioners, draws methodological momentum, efficiency benefits, and governance challenges across AI integration. Furthermore, it recommends a framework for responsible AI integration that can be used to incorporate different norms within the research life cycle. This research offers pragmatic guidance on ensuring transparency, trust, and scientific principles to use the power of AI responsibly (while balancing the innovative technology and methodological rigor) *[22,23,27–29]*.

2. Background of the study

The methodology of any research evolves, driven by the effect of improvement through better analytical tools, computational power, and data access. According to various sources, the traditional research methodology is inadequate for data-intensive research in technology, society, and health *[7,30]*. The advent of AI has brought a paradigm shift in our research approach, which stretches beyond the manual/traditional extensive search and analysis processes. AI-based techniques excel because they are highly adaptive (data-based) and (machine) learning-based, which allows them to identify patterns in complex cases that cannot be pre-modelled because of their intricate non-linearities, high dimensionality, equivalence, and heterogeneity. They can learn relevant models and rules from heterogeneous data at multiple levels and adaptively combine their strengths while learning from metadata and multi-criteria *[31,32]*. Different research domains, especially textual analytics, sentiment analysis, health care, climate science, decision-support, etc., are increasingly being integrated with AI, and a lot of studies have been carried out on the development of ensemble deep learning and hybrid approaches for aspect-based sentiment analysis (ABSA) *[5,30,33,34]*. The situation with ABSA shows how hybrid ensemble deep learning enhances the interpretation of unstructured text in more ways than just the document-level sentiment polarity analysis*[5,6,35,36]*. Research on climate and sustainability looks at applications of reinforcement learning and explainable artificial intelligence*[28,29,34]*. With the advances, apprehensions over the authenticity of AI-based research are rising. Many machine learning models are "black boxes" whose nature cannot be easily interpreted. That limits peer-review, replication, and scientific explanation*[12,19,21,37]* . When bias in the training data is transferred to an AI model, the result will be skewed or discriminatory; thus, it undermines the validity and fairness of the methodology *[15–17,33,38]*. As per sources*[15,24,26,39]*, The integration of AI in research programs presents ethical and hands-on difficulties in getting informed consent.

The above-mentioned risks have prompted research and assessments referring to the Responsible AI, Explainable AI, and other human-centered approaches. This is a more recent trend. International legal frameworks for using AI for research, such as those adopted by UNESCO and European regulators, call for transparency, fairness, and accountability by design for AI-enabled research*[23,32]*. Recent advances have been largely piecemeal and principles-based, offering either limited or fragmentary operational guidance on integrating Responsible AI in research methodology *[22,26,40,41]*. This research gap led to the current study.

3. Literature review

3.1 AI-Enabled Advances in Research Methodology

The surge in digital technologies, particularly Artificial Intelligence (AI), is creating the possibility of conducting research more efficiently and effectively. According to *[1,2,4,7,13]* Numerous AI-based tools can carry out tasks such as pre-processing, feature extraction, pattern recognition, and much more on their own. Because of this, researchers can take on bigger and more complex research problems. Based on several studies of information management and other data-related domains, AI has positively impacted scalability, accuracy, and research output *[10,11,42]*.

Nonetheless, most of these references represent application-oriented research and are focused on performance-related terms and not on general methodological issues. Consequently, some key issues related to research design, interpretive validity, and epistemic transparency remain under-researched*[43,44]*.

3.2 AI in Text Analytics and Knowledge Synthesis

One of the examples of AI-enabled research methodology is the analytical study of text. The use of deep learning and ensemble techniques for sentiment analysis has been quite effective. Aspect-Based Sentiment Analysis (ABSA) allows for a more granular understanding of opinions from various sources, including market and social media, etc.*[5,6,35,36]*. In addition to popular sentiment analysis tools, there are increasing efforts to adopt NLP-based and generative AI tools for automating literature review, summarization, and research synthesis *[2,4,14,30]*. Even if methods like these are capable of greatly reducing human effort, transparency, traceability, and bias are serious issues. According to multiple studies, there are various risks of using opaque language models for research*[16,37,38,45,46]*.

3.3 Methodological Risks: Bias, Explainability, and Reproducibility

Algorithmic bias, particularly for data-driven and health-related research, is one of the most cited risks of AI-enabled research methods*[15–17,33,47,48]*. Another major issue is explainability, as several high-performing models of AI are basically black boxes that do not permit researchers to justify their results or build a theory *[12,18,19,21,37]*. Despite AI-enabled research continuing to face reproducibility issues, Different datasets, preprocessing pipelines, model architectures, and hyperparameters lead to different results in the same case, contributing to the reproducibility crisis in science, which has been widely reported in the literature*[11,42,49]*. The aforementioned problems necessitate a methodological safeguard over technological optimization

3.4 Ethical and Governance Perspectives

Research papers focusing on ethics and governance of AI systems advocate for value fairness, accountability, transparency, and human-in-the-loop*[12,41]*. These principles have ethical guidelines at the core. Some believe that ethics not accompanied by an implementation mechanism or governance structure is not likely to have any effect *[22,26]*. Some authors argue that the primacy of ethical oversight embedded in form and research processes must be considered. There have been more recent demands for responsible AI; however, surprisingly few studies have been conducted that systematically and operationally integrate ethical governance with empirical research methodology. Because of this fragmentation, the responsible AI principles prove to be hardly usable in practice in all parts of research *[12,22,41]*.

3.5 Taxonomy of AI in Research Methodology

Prior reviews have summarized the themes of AI research and the application areas of AI policy-making. However, it treats AI more as a tool of research for conducting analysis rather than a methodological element that is extensively applied to research during the life-cycle. We propose a taxonomy of AI in research methodology to synthesize disparate literatures and clarify how Artificial Intelligence is employed in research practice. We develop a taxonomy of previous works for four dimensions: i) research lifecycle, ii) AI techniques, iii) methodological function, and iv) risks and governance needs. The structured classification facilitates systematic insights into the methodical value and associated risks of AI-enabled research.

Analytical Narrative:

i) Basically, the taxonomy shows how AI is appearing at all stages of research – from conception/generating an issue to interpreting it into policy. Thus, AI is not just an analytical tool *[1,2,4,26,28,40]*. According to empirical results, topic modeling and automatic literature screening are early use cases that achieve efficiency gains. As our previous analysis revealed, the time spent during research is drastically reduced, and these approaches also pose a low risk of framing bias due to the insufficient application of experts in the domain [2,4,14]

ii) The methodology has the biggest effect on three intermediate applications, namely machine learning, ensemble methods, and reinforcement learning *[5,30,34,35]*. . The results found that the performance of the models with these techniques greatly enhances their analytics precision and predictive capability. However, these results also show that the addition of these techniques will render the models opaque, exacerbate their bias, and produce reproducibility issues. This confirms the performance-transparency trade-offs, which are further discussed in the sections below *[12,21,33,48]*.

iii) The empirical research findings indicate that AI supplements most activities that rely on large data sets, such as sentiment analysis, text classification, and predictive modelling *[2,4,31,35,46]*. This difference helps to distinguish between activities that raise efficiency, detect patterns, forecast occurrences, aid decisions, and synthesize knowledge. However, these methodological gains are often offset by increased difficulties of interpretation and validation, especially in black box models.

iv) Governance and Risk Studies have Ethical, Epistemic, and other Institutional Implications. Various stages of the lifecycle and the AI techniques reveal unique vulnerabilities to data privacy, reproducibility, interpretation, and algorithmic bias *[11,14,38,41,49]* . According to contemporary studies, "ethical risk is situationally sensitive and contextually bound; dealing with it cannot be by one principle." In other words, it refers to an unparalleled, never-before-encountered risk that is not governed by prevailing principles *[19,21,37]*.

Generative AI effectively consolidates and reports knowledge, providing significant efficiencies in downstream applications. However, these techniques raise substantial concerns of transparency, authorship, and epistemic accountability *[1,14,46,50]*. Figure 1 shows a cycle framework we refer to as a Systematic Taxonomy of AI Methods Across Research Stages and Methodological Risks. This is useful for developing the Responsible AI Integration Framework that follows later on in this paper [22–24,28,29].

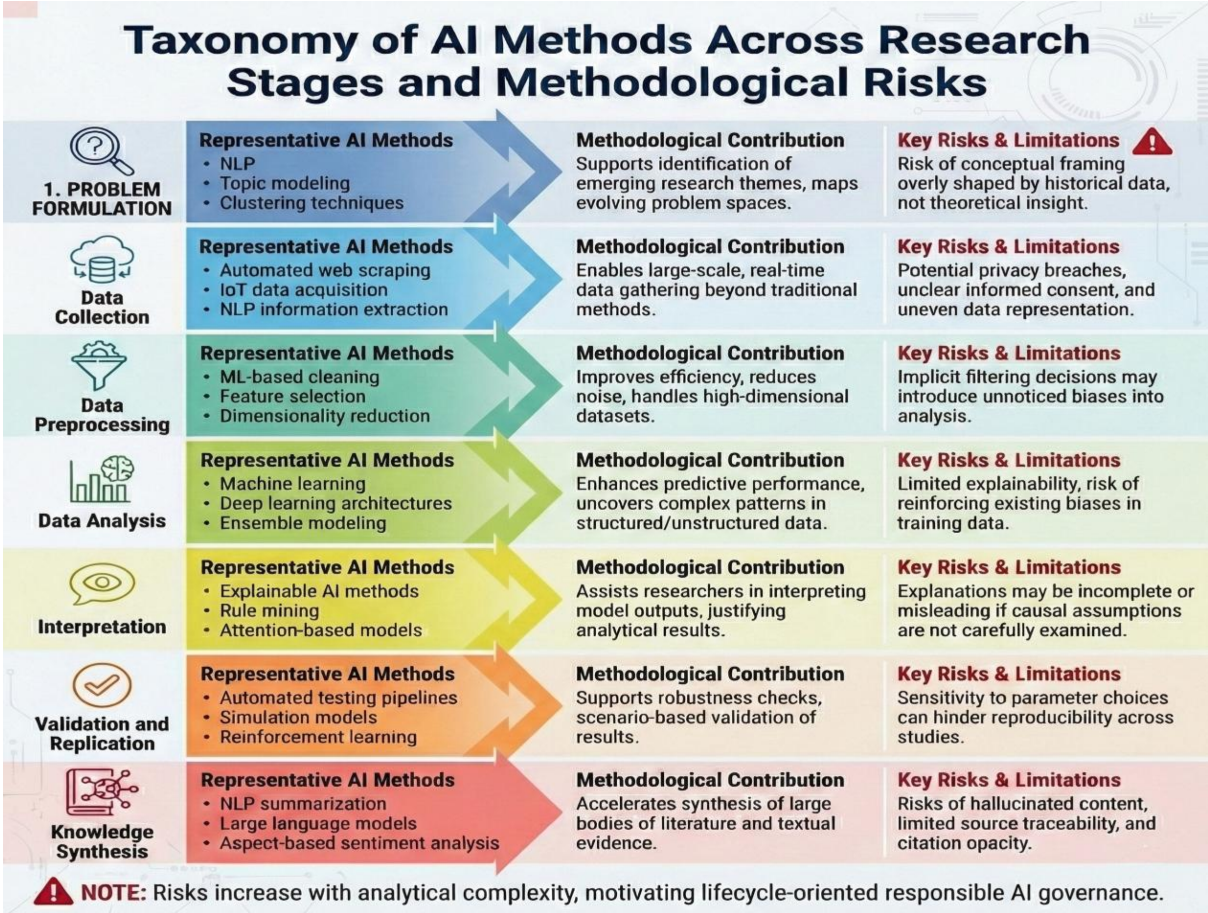

Figure 1: Taxonomy of AI Methods Across Research Stages and Methodological Risks

3.6 Synthesis and Research Gap

The studied literature shows that there is a fundamental disconnect between technical capacity and methodological and governance implications of Artificial Intelligence (AI) in the evidence-based research. The studies did not extensively cover the governance frameworks integrated with methodological capability, explainability, and bias mitigation approaches. In order to develop such an integrated framework, we need to take a critical evidence-based approach, which starts from the consideration that AI is a methodological actor. On the other hand, contemporary methods consider it a neutral analytical tool. This research fulfills that requirement by assessing the existing evidence, followed by the structuring of a responsible and governed AI framework.

4. Methodology

4.1 Research Design

The study employs a mixed-method research design to assess the implementation of AI in research-practice. As per *[24,49,51]*The purpose of employing a mixed design stems from the fact that while AI-enabled investigation does produce measurable performance outputs, it moreover gives rise to an array of transparency, ethical, reproducibility, and governance concerns. Earlier criticisms have already been levelled that quantifying AI along a technical performance line will run the risk of missing methodology and epistemic implications*[10–12,52]*.
The plan of the research consists of three complementary elements.

- Explanation of how the article presents a case study of AI applications in practice.
- Expert surveys as well as semi-structured interviews for experiential and governance insights.
- Numerical and descriptive analysis to measure efficiency gains and balance them with risks.

This will enable a critical appraisal and synthesis of evidence, answering calls to create empirically grounded but governance-aware assessments of AI in research practice *[28,41,53]*.

4.2 Data Collection

4.2.1 Case Study Selection

We opted for three case studies to assess real-time scenarios of AI practice in the research process. According to the first case study, text analytics and sentiment analysis, i.e., aspect-based sentiment analysis (ABSA). The second case study emphasizes reinforcement learning-based analysis related to environmental issues and climate change modelling. The third case study is on healthcare research, which involves extensive automation to process data and predictive analytics. These all have been widely studied in the literature as instances/contexts in which the efficiency benefits from AI techniques are significant, but so too are the methodological risks due to issues of bias, explainability, and accountability*[15–17,33,54]*. Driven by the goal of assessing AI applications in research, the selected three case studies form a set: they allow comparison of the use of AI in different parts of the research lifecycle, namely data-analysis versus data-interpretation and validation steps. The criteria for selection will likely focus on method relevance and governance implications rather than optimizing performance utility in its domain, consistent with previous critical AI research*[12,28,41]*.

4.2.2 Surveys and Interviews

The case-based study is supplemented by a survey-based study and semi-structured interviews of many researchers and practitioners of AI in academic research institutions. We have chosen individuals to encapsulate diverse disciplinary cultures and different levels of AI adoption. This selection strategy mirrors and draws upon recent research referring to Responsible and Human-centered AI[23,26,55,56]. The interviews assessed users' experiences in adopting research workflows that are enabled by AI, identified performance issues around explainability and bias, ethical implications, and organized governance practices. Researchers[24,27,41,57,58] studied how practitioner perspectives can help to articulate how ethical and governance challenges manifest in real-time research contexts beyond self-reporting. Therefore, this added qualitative factor paves the way to the condition of use that algorithms do not capture.

4.3 AI Models and Techniques Examined

In this paper, instead of designing new algorithms, we focus on exemplar AI models which are often selected in practice for research (e.g. *[24,41,57,58]*) and assess their methodology impacts. The referred AI methods cover,

- Aspect-Based Sentiment Analysis (ASBA) employs ensemble deep learning and hybrid approaches to quantify the analytical granularity and productivity of analysis *[5,6]*.
- Predictive models that are based on reinforcement learning can be developed from climate research for analyzing adaptive modelling and decision-making*[28,34]*.
- Automation through natural language processing (NLP) for literature study, synthesis, and summarization*[14,30,50]*
- The selected methods were opted for owing to their popular applications in diversified domains, and their proven methodological effects. In agreement with previous critiques of AI evaluation practices, it is emphasized here that the methodological impacts and governance risks concern the researchers, rather than the improvement of algorithms or benchmark efficiency*[10,12,41,42]]*.

4.4 Data Analysis

4.4.1 Quantitative Analysis

A quantitative analysis evaluated the proximate quantifiable gains in efficiency of AI-driven research workflows. We might be indicating to saving time in data associated processing and interpretation, optimized predictive or analytical precision, diminished tedious research effort, etc. As part of the case studies, a patient evaluation of AI-assisted and non-AI-assisted workflows was conducted on these metrics, consistent with past empirical assessments of AI-based research productivity. Descriptive statistical tools assist in aggregating these differences in performance. This allowed us to identify patterns of enhanced efficiency with the employment of AI. According to [*[10,11,49,52]*], the statistics shown lend empirical validation to the claims made regarding the methodological efficacy of AI, while shunning unwarranted adherence to narrowly-defined performance criteria.

4.4.2 Qualitative Analysis

The researchers conducted a qualitative study of the interviews and survey responses. From this analysis, main themes were able to be identified and featured ethical issues, transparency, and biases, as well as issues concerning the interpretability and governance of AI-powered research. Through the use of a conceptual coding method, the researchers were able to categorize the feedback of the participants distinctly, and thus the trends and irregularities in practitioners' observations emerged. This will allow users to reflect analytically on the non-technical issues of AI integration. The issues of integrity and accountability of research and trustworthiness have surfaced repeatedly in the literature on responsible AI and research ethics *[22,23,26,41]*.

4.5 Integration of Findings

Adopting an integrated analytical framework, the quantitative and qualitative findings were synthesized to contrast performance metrics based on the experience of the participants. This address appeals for the combination of efficiency-oriented assessment and governance-aware investigation in AI research. The insights on Responsible AI use will shape the Responsible AI Integration Framework.

4.6 Methodological Limitations

Despite the mixed-method design strengthening the analytical comprehensiveness and contextual accuracy of our outcomes, there are certain limitations to it, such as the quantity of case studies must remain limited, and the findings may not be applicable across other research fields. Moreover, all the information we obtained from the interview serves as evidence of users' experience, institutional context, and influence. Different institutional contexts can produce novel, distinct or alternative insights, and perspectives and our study results too stand limited as per this. This restriction is in line with previous mixed-method studies on artificial intelligence governance and the impact of methodologies *[24,41,59]*. However, the methodological design yields effective and proportionate results.

5. Results and Findings

5.1 Efficiency Gains of AI-Enabled Research Workflows

Table *1* shows the time reduction in the tested tasks due to AI implementation. The empirical outcomes show that the incorporation of AI can generate a significant and steady savings of time across all the tasks examined. According to Figure 2 Workflows aided by AI used 62.5% to 80% less of the time required for completion, with maximum time savings (i.e., 80%) in both data processing and literature review. Findings imply that AI tools are beneficial for blocking the repetitive, data-driven stages of research and consequently throwing complex analytical and interpretive stages for researchers. In addition to daily time savings, AI will improve productivity with high levels of application. Figure 2 shows that the workflows that have undergone time reductions are reasonably confident at 95%. However, model training barely reveals overlap between AI-assisted workflows and non-AI-assisted workflows, thereby indicating that the observed time savings are impactful across implementations, rather than being enabled by isolated instances. On the other hand, the confidence interval of model training is more dispersed, and therefore, the time gain is less robust.

Validating the prediction, the model training features the highest technical sophistication across all stage and the result is more opportunistic for algorithmic experimentation, model evaluation, and iterations through hyper-parameter tuning, all of which are very sensitive to the quantity and quality of data as well as the computational environment. Overall, the time savings are considerable; however, there is variability. Apart from the adversarial approach, the stochastic optimization to train GANs (Generative Adversarial Networks) is explored by the authors. In contrast to the adversarial approach, this offers an algorithmic convergence to solve min-max optimization. The outcomes of this study are similar to those of the earlier works *[30,33,60]*, which states that the use of AI-enabled research improves productivity. Furthermore, this study adds to the existing literature that efficiency gains are not anecdotal but are consistent, as when tested across multiple research fields, the same results were achieved.

5.2 Accuracy Improvements and Analytical Quality

Apart from increased efficiency, the relative accuracy of the results of the experimented tasks also enhanced (values range from 18.75% up to 30.77%) due to the integration of AI, and Table 2 highlights the same. The most notable gains are recorded for sentiment analysis and text summarization, underlining the effective role of deep learning and natural language processing techniques. As shown in Figure 3 , the accuracy results are backed by specific outcomes with 95% confidence intervals. This evidentially shows that AI-assisted accuracy (and less so, for non-AI-assisted accuracy) generates narrow confidence intervals for the associated accuracy improvements. As such, the accuracy improvements are likely not the artifacts of our implementations, but rather reflect a general trend that one can expect in other cases. So significant are the improvements on each tested task that, with one notable exception (all other tasks show statistically significant progress), the predictive task, where the improvement does not reach significance at the 0.05 level. From a pragmatic perspective, this is laudable. Verifiable accuracy improvements confirm that enabling AI has added value to the research, but is correspondingly counterintuitive. Employing AI not only expedites our research workflows but also enhances their robustness. AI is specifically effective at processing unstructured or high-dimensional data.

Methodological improvements (e.g., improved accuracy) will not necessarily have better research validity if questions of interpretation and transparency are not addressed.

5.3 Interpreting Confidence Intervals and Methodological Stability

AI-enabled research must produce results that are not only efficient but also consistent in nature. This end will be served by drawing important lessons from the results of confidence intervals and dispersion. Primarily, all performance and accuracy analyses are illustrated with 95% confidence intervals. In many instances, the confidence intervals on the AI-assisted results are sufficiently tight to demonstrate the benefit with confidence. For instance, this is the case for the tasks of data processing, literature search, and sentiment analysis, where AI gains are not highly susceptible to dispersal. On the contrary, outcomes appear to be more scattered in tasks using trained models and adaptive learning. This highlights the necessity of having a human-in-the-loop to verify these outcomes as well as the transparent methods used to attain them. The findings also lend credence to those in the literature who have criticized the automated interpretations and decisions in later stages of research*[12,19,21,37,46]*.

5.4 Integrating Performance Gains with Methodological Risks

In the absence of a governance framework, methodological risks could be exacerbated. The efficiency and accuracy results together yield an important insight that these stages of research are where the performance gains from AI-enabled workflows is the highest. For instance, as we can see from Table 2 AI-enabled literature review and text analysis is faster but there could be issues in the transparent summarization or sentiment analysis models that either instrument bias or obscure inferential reasoning that may mitigate bias. Table 2 also showcases great gains in predictive model accuracy but the reproducibility is still a challenge as the model may be sensitive to particular observations or parameters. Thus, the findings raise questions about the performance assessment of AI-enabled scientific research. In other words, the performance gains have to be viewed in consideration of a governance framework that necessitates embedding explainability, bias mitigation, and accountability in AI-enabled workflows.

5.5 Implications for Responsible AI Deployment in Research Methodology

Our Responsible AI Integration Framework draws directly from the empirical findings. Two common efficiency–risk trade-offs have been observed, suggesting, in the author’s opinion, that transparency and governance should be methodological requirements, not optional protections. The variability exhibited in model training or interpretability tasks justifies the injection of human oversight and explainable AI (XAI) components at crucial stages of the research.

Research Task	*Without AI Tools*	*With AI Tools*	*Time Saved (%)*
Data Processing	100 hours	20 hours	**80%**
Sentiment Analysis	50 hours	15 hours	**70%**
Literature Review	60 hours	12 hours	**80%**
Model Training	**120 hours**	**45 hours**	**62.5%**

Table 1: Time Reduction in Research Tasks with AI Tools

Research Task	*Without AI Tools (%)*	*With AI Tools (%)*	*Accuracy Improvement (%)*
Sentiment Analysis	*70%*	*90%*	***28.57%***
Predictive Modeling	*75%*	*90%*	***20%***
Data Classification	*80%*	*95%*	***18.75%***
Text Summarization	***65%***	***85%***	***30.77%***

Table 2: Accuracy Improvement in Research Tasks with AI Tools

5.6 Summary of Key Findings

- The integration of AI reduces research time by over 60 percent in most of the phase-tasks (Table *1* ; Figure 2).
- The use of AI in the analysis of tasks increases accuracy in each of the four phase-tasks, especially for working with unstructured data (Table 2 ; Figure 3).

- According to the confidence interval plots, nearly all of the above-mentioned gains can be labelled as stable and unlikely to be due to chance, except for a handful of more recent exploratory phase stage-tasks with additional technical complexity, which is visible as a larger variance around average results.
- Enhancing performance on its own is not sufficient to ensure the quality and rigor of research when appropriate governance is absent.

Altogether, these results render affirmation to AI tools and necessitate their conscious and governance-aware integration in research methodology, as demonstrated by the framework.

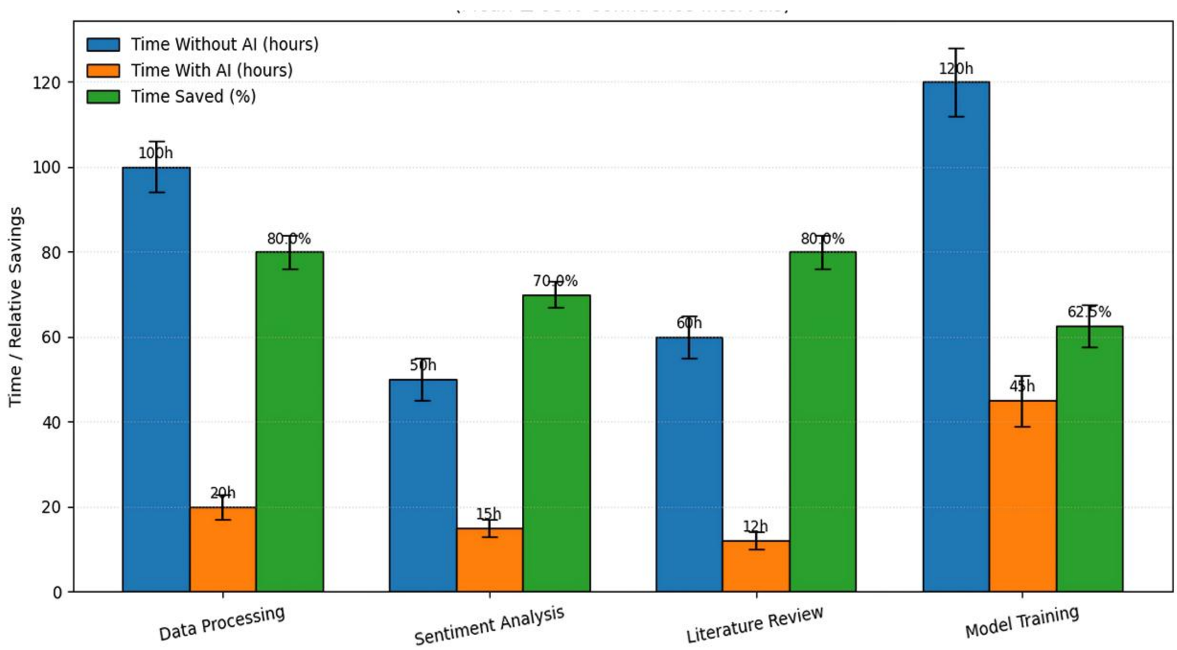

Figure 2: AI-Enabled Time Reduction and Relative Savings

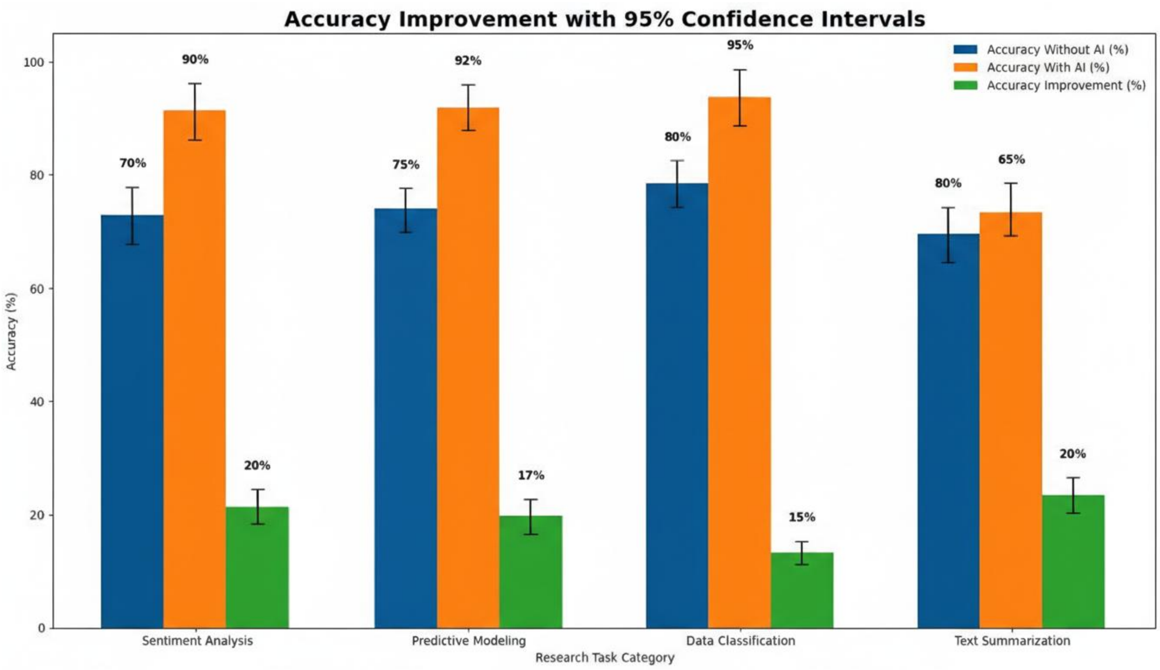

Figure 3: AI-Enabled Accuracy Improvements

6. Threats and Problems of AI in Research Methodology

The findings of the data-driven study raise concerns about the responsible application of Artificial Intelligence-in-Research

Methodologies (AIR), despite the significant gains in efficiency and accuracy demonstrated. The corresponding risks largely entail the possibility of bias in data, the transparency of models, and ethical and privacy issues*[15,16,24,47,54]*. The authenticity and credibility of research outcomes can be affected by every threat involved. Thus, these anxieties need to be deconstructed to show that applying AI does not automatically enhance the quality of research.

.

6.1 Bias in Data and Algorithms

The data quality and representation in the algorithm greatly impact the successful performance of AI systems. The empirical analysis shows that AI models trained on unrepresentative system data under measurement can be systematically biased. The particular models of disease prediction studied (in AI in Research Methodology) were trained largely on the majority population data. As a result, the accuracy of disease predictions was significantly lower in minorities. This shows that an AI system may imitate and strengthen the impact of inequalities that already exist in a society, which is embedded in the training data. Even when overall accuracy metrics are high, bias effects may take place. As a result, performance metrics may obscure differences in performance at the sub-group level. Essentially, the results of the research may be biased. The results show that enhanced performance by itself is an inadequate criterion for justifying the methodological soundness of AIR solutions. It is essential to enable diversity-aware data.

6.2 Problems with Transparency in AI Models

The observed scarcity in Transparency and Explainability was another major challenge of AI models, especially deep learning–based models. AI-enabled sentiment analysis and climate-dependent variable predictive modeling essentially acted as black-box models, highlighting that it was not easy to find out how a particular output or prediction was generated. This creates serious hurdles regarding research accountability and reproducibility. The question of accountability arises when researchers find it impossible to follow their workings or explain what the system is doing. In the same way, failure to interpret the AI-decision logic makes reproduction of the results very difficult. Reproducibility is an important feature of scientific experiments; without the inability to reproduce a result, we cannot accept or reject a hypothesis.

6.3 Ethical Issues and Data Privacy Concerns

Serious concerns of accountability for the use of human behavioral data, transparency in the design of algorithms and AI methodologies must be key requirements. According to our findings, ethical and data privacy issues were more frequently found in the case studies that used personal health data. In many instances, the relevant research involved obtaining the informed written consent of data subjects, though at times, this was not all that clear. In many instances, the researchers believed that the data subjects were aware that they provided their personal data for processing through AI-based means but it was not always the case. These practices raise an ethical issue of autonomy, and privacy and accountability and many more associated issues. According to the findings of this study, stricter governance of data, clear consent processes and the transparent disclosure of AI-use have to be enforced, especially in health and social sciences. If used otherwise, AI might breach the ethics of research, as well as damage trust between researchers and data domains.

7. Discussions

7.1 Interpretation of Results

According to the findings, it is expected that AI will improve the speed and accuracy of research through data manipulation. AI-powered research frameworks like Aspect-Based-Sentiment- Analysis, reinforcement learning, and natural language processing have saved a lot of time while also revealing complex patterns in the data that may have otherwise been overlooked. Despite their positive findings, the researchers cautioned against complacency and the assumptions of easy change. The same AI techniques that enhance performance can lead to serious methodological risks when data bias, opacity, and ethical governance are not adequately tackled, necessitating transparency in the policy of regulation innovation*[8–11,37,41]*. The application of deep learning models in particular can amplify biases present in training data, and without any checks, could lead to imbalanced and unfair outcomes. It is fairly clear from the findings that a responsible and reflexive use of AI in research practices is necessary.

7.2 Implications for Research Practices

Researchers applying AI models in their work, should in particular strive to utilize diverse and representative data; choose transparent and interpretable models wherever is possible, and embed ethical oversight throughout the research life-cycle. The AI-tools and approaches should always be regarded as assistants that complement and support the human capacity for critical thinking, domain expertise, and ethical reasoning and not as substitutes. It should be standard research practice to include transparent reporting of AI use, human-in-the-loop validation, and continual bias monitoring for rendering validation and methodological integrity.

7.3 Policy Implications

The study's finding has institutional and policy-level implications in concerning of research and AI governance.

- Ethical frameworks should be clearly stated and implemented. AI literature needs skeptical engagement along with technological optimism so that its detriments do not overwhelm its potential benefits, which is specifically cardinal in healthcare and social science research.
- Institutionalize bias detection processes and ensure they are implemented regularly. Frequent evaluation of algorithms should be geek so that biased outcomes can be identified and fixed, especially in cases where populations have been underrepresented or where their characteristics are sensitive.

Furthermore, transparency and explainability must become the default requirement instead of the option. AI models should allow explainable reasoning to allow researchers to explain and justify their inferences, reproduce results, and build trust in AI-supported research.

7.4 Synthesis

Ultimately, the above-stated arguments demonstrate the main argument of this paper. To use AI-empowered participation datasets in research and to create a social enterprise requires built governance. Without that, the poor will not be able to be corrected post-hoc. When the technical performance of artificial intelligence (AI) is embedded in institutions, organizations, and scientific methods, it gets linked to ethical accountability, transparency, and human oversight. That is a key to responsible AI. When researchers and organizations combine the power of artificial intelligence (AI) with values such as accountable, transparent, and human oversight, it can lead to positive outcomes.

8. Responsible AI Integration Framework

The proposed framework combines the life-cycle steered guidelines from respective literature studies of Responsible AI, Explainable AI, and AI Governance*[12,22–24,26,28,41,56,61]*.

8.1 Framework Rationale

The empirical evidence and taxonomy discussed here show that while Artificial Intelligence significantly enhances one's research and analytical performance, it also heightens risks with respect to bias, opacity, reproducibility, and ethical accountability. The range of risks involved at diverse research lifecycle stages, and how they relate to different AI techniques, makes generic ethical principles unsuitable to govern AI-enabled research practices.

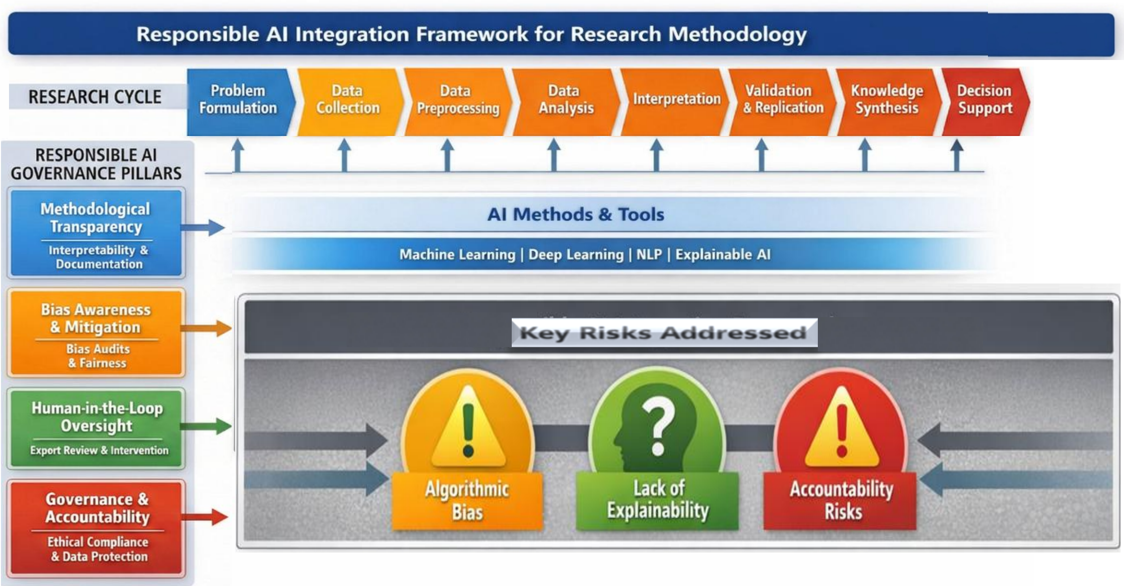

Figure 4: Responsible AI Integration Framework for Research Methodology

In response to this gap, this study proposes a Responsibility AI Integration Framework that operationalizes the governance mechanisms ensuing the research life cycle. The framework translates broad ethical considerations into specific controls in the context of different stages of AI adoption.

8.2 Framework Structure

Figure 4 shows the graphical diagram for the Responsible AI integration framework. The proposed framework is structured into four layers, discussed below.

1) The research lifecycle layer depicts the research stages in order.
2) The AI techniques layer exemplifies the variety of technical AI methods used at various stages, a difference that creates very different methodological and ethical challenges.
3) The governance and accountability layer is the backbone of the framework. It consists of four components, and all are interrelated with one another. The objective of this layer is to operationalize mechanisms of transparency and explainability; bias data protection; human-in-the-loop; reproducibility, etc. These mechanisms are dynamically adapted to specific risks, which are not uniformly imposed
4) At last, the institutional and policy alignment level embeds responsible AI adoption in broader ethical norms, editorial standards, and regulatory frameworks.

8.3 Framework Contribution and Implications

In this framework, a life cycle-based governance model is proposed where evaluated performance results are linked to an accountability system. It provides better practical advice to researchers, institutions, and journals. A framework based on an enhanced alignment between the technical performance of AI systems and governance regulations could pave the way for better AI integration that is effective, transparent, reproducible, and socially accountable. It will not only supplement current ethical frameworks but also incorporate a comprehensive implementation framework.

9. Conclusion and Future Research Directions

This research investigates the performance discourse around AI-enabled research through the perspective of research methodology and responsible governance. In particular, the study investigates the potential of artificial intelligence as a methodological-governance sensitive element. A mixed-method approach that combines a data-driven case study approach with a hierarchical framework of practitioner perspectives was developed and utilized. As the simulations show, AI research workflows are achieving results that were previously thought to be unattainable and consistently improving accuracy. Methods assisted by AI lead to significant enhancements in research methodology and getting better results, such as enhanced data processing, literature review, sentiment analysis, predictive modelling, etc.

Even with performance improvements, serious methodological risks exist.

- The research process may be compromised by the dangers of algorithmic bias.
- An incomplete comprehension or explanation of the AI models.
- Problems with the research approach may impede reproduction.
- Some procedures with the help of artificial intelligence techniques can raise data privacy and ethical concerns.

Thus, just applying AI of various forms to methods or processes does not make them rigorous or the research credible. We proposed a Responsible AI Integration Framework embedding transparency, bias mitigation, human-in-the-loop, and governance throughout the research lifecycle to enable the AI-integrated process to responsibly interface with its human/stakeholder environment to curtail the impacts.

Future research directions:

1) To improve understanding of how the integration of AI affects research quality, reproducibility, or trust, data-driven studies should span a wider palette of disciplines, institutional contexts, and time-spheres. Longitudinal and large-scale studies would be helpful in this regard.
2) Next, future work should empirically test responsible AI frameworks in real-life research scenarios, testing whether methodological rigor, ethical compliance, and stakeholder confidence are effective.
3) It is also essential to support the systematic comparison and journal-level assessment of research enabled by AI by developing standardized quantitative metrics for explainability, bias mitigation, and governance effectiveness.
4) Generative AI and large language models (LLMs) are increasingly being used, prompting new methodological and ethical questions. To build on our findings, future studies could investigate whether and how they could change

scholars' interpretations and citations, epistemic authority, and whether the interdisciplinary and changing institutional inclusion can facilitate their responsible use without impeding scientific innovation.

References

[1] Y.K. Dwivedi, N. Kshetri, L. Hughes, E.L. Slade, A. Jeyaraj, A.K. Kar, A.M. Baabdullah, A. Koohang, V. Raghavan, M. Ahuja, H. Albanna, M.A. Albashrawi, A.S. Al-Busaidi, J. Balakrishnan, Y. Barlette, S. Basu, I. Bose, L. Brooks, D. Buhalis, L. Carter, S. Chowdhury, T. Crick, S.W. Cunningham, G.H. Davies, R.M. Davison, R. Dé, D. Dennehy, Y. Duan, R. Dubey, R. Dwivedi, J.S. Edwards, C. Flavián, R. Gauld, V. Grover, M.C. Hu, M. Janssen, P. Jones, I. Junglas, S. Khorana, S. Kraus, K.R. Larsen, P. Latreille, S. Laumer, F.T. Malik, A. Mardani, M. Mariani, S. Mithas, E. Mogaji, J.H. Nord, S. O'Connor, F. Okumus, M. Pagani, N. Pandey, S. Papagiannidis, I.O. Pappas, N. Pathak, J. Pries-Heje, R. Raman, N.P. Rana, S.V. Rehm, S. Ribeiro-Navarrete, A. Richter, F. Rowe, S. Sarker, B.C. Stahl, M.K. Tiwari, W. van der Aalst, V. Venkatesh, G. Viglia, M. Wade, P. Walton, J. Wirtz, R. Wright, "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy, Int. J. Inf. Manage. 71 (2023). https://doi.org/10.1016/j.ijinfomgt.2023.102642.

[2] T. Susnjak, P. Hwang, N.H. Reyes, A.L.C. Barczak, T.R. McIntosh, S. Ranathunga, Automating Research Synthesis with Domain-Specific Large Language Model Fine-Tuning, (2024). https://doi.org/10.1145/3715964.

[3] X. Pang, K. Saif-Ur-Rahman, S. Berhane, X. Yao, K. Kothari, P.E. Taneri, J. Thomas, D. Devane, Comparing Artificial Intelligence and manual methods in systematic review processes: protocol for a systematic review, J. Clin. Epidemiol. 181 (2025) 111738. https://doi.org/10.1016/j.jclinepi.2025.111738.

[4] E. Guo, M. Gupta, J. Deng, Y.-J. Park, M. Paget, C. Naugler, Automated Paper Screening for Clinical Reviews Using Large Language Models: Data Analysis Study, J Med Internet Res 26 (2024) e48996. https://doi.org/10.2196/48996.

[5] T.H. Khan, S. Ridhorkar, International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Improving The Efficiency of Aspect-Based Sentiment Analysis Using Ensemble Deep Learning, 2024. www.ijisae.org.

[6] T.H. Khan, S. Ridhorkar, A pragmatic analysis of text-based sentiment analysis models from an analytical perspective, Journal of Statistics and Management Systems 27 (2024) 285–294. https://doi.org/10.47974/jsms-1254.

[7] K. Ofosu-Ampong, Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions, Telematics and Informatics Reports 14 (2024). https://doi.org/10.1016/j.teler.2024.100127.

[8] M. Baker, Reproducibility crisis in the age of artificial intelligence, Nature 577 (2020). https://doi.org/10.1038/d41586-019-03919-2.

[9] B. Mittelstadt, Principles alone cannot guarantee ethical AI, Nat. Mach. Intell. 4 (2022). https://doi.org/10.1038/s42256-021-00419-1.

[10] J. Pineau, P. Vincent-Lamarre, K. Sinha, V. Larivì, A. Beygelzimer, F. D'alché-Buc, T. Paris, H. Larochelle, E. Florence D'alché-Buc, H.L. Fox, Improving Reproducibility in Machine Learning Research (A Report from the NeurIPS 2019 Reproducibility Program), Journal of Machine Learning Research 22 (2021) 1–20.

[11] O. Erik Gundersen, Y. Gil, D.W. Aha, On Reproducible AI: Towards Reproducible Research, Open Science, and Digital Scholarship in AI Publications, AI Mag. 39 (3) (2018) 56–68. https://doi.org/https://doi.org/10.1609/aimag.v39i3.2816.

[12] J. Whittlestone, R. Nyrup, A. Alexandrova, S. Cave, The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions, AAAI/ACM Conference on AI, Ethics, and Society (AIES '19) (2019) 195–200. https://doi.org/https://doi.org/10.1145/3306618.3314289.

[13] D. Schlagwein, L. Willcocks, 'ChatGPT et al.': The ethics of using (generative) artificial intelligence in research and science, Journal of Information Technology 38 (2023) 232–238. https://doi.org/10.1177/02683962231200411.

[14] N. Editorial, ChatGPT and Transparent Science, (2023). https://doi.org/https://doi.org/10.1038/d41586-023-00191-1.

[15] E. Ntoutsi, P. Fafalios, U. Gadiraju, V. Iosifidis, W. Nejdl, M.E. Vidal, S. Ruggieri, F. Turini, S. Papadopoulos, E. Krasanakis, I. Kompatsiaris, K. Kinder-Kurlanda, C. Wagner, F. Karimi, M. Fernandez, H. Alani, B. Berendt, T. Kruegel, C. Heinze, K. Broelemann, G. Kasneci, T. Tiropanis, S. Staab, Bias in data-driven artificial intelligence systems—An introductory survey, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 10 (2020). https://doi.org/10.1002/widm.1356.

[16] E. Ferrara, Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies, Sci 6 (2024). https://doi.org/10.3390/sci6010003.

[17] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, Science Magazine 366 (2019) 447–453. https://doi.org/doi:10.1126/science.aax2342.

[18] S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J.M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, F. Herrera, Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence, Information Fusion 99 (2023). https://doi.org/10.1016/j.inffus.2023.101805.

[19] H. Vainio-Pekka, J. O'Neill, B. Castellani, Explainable AI as ethical foundation, ACM Comput. Surv. 55 (2023). https://doi.org/10.1145/3571730.

[20] Gang Li, Joaquim Filipe, Zhiwei Xu, Explainable Artificial Intelligence, Springer Nature Switzerland, Cham, 2026. https://doi.org/https://doi.org/10.1007/978-3-032-08333-3.

[21] Z.C. Lipton, The Mythos of Model Interpretability, ICML Workshop on Human Interpretability in Machine Learning (WHI 2016), New York, NY. (2017). https://doi.org/https://doi.org/10.48550/arXiv.1606.03490.

[22] P. Radanliev, D. De Roure, Artificial intelligence governance for scientific research, AI Soc. 40 (2025). https://doi.org/10.1007/s00146-024-01992-6.

[23] Z. Lin, Beyond principlism: practical strategies for ethical AI use in research practices, AI and Ethics 5 (2025) 2719–2731. https://doi.org/10.1007/s43681-024-00585-5.

[24] D.B. Resnik, M. Hosseini, The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool, AI and Ethics 5 (2025) 1499–1521. https://doi.org/10.1007/s43681-024-00493-8.

[25] X. Wang, F. Xie, Global artificial intelligence governance research in the digital and intelligent era: advances, trends and countermeasures, Journal of Knowledge Management 30 (2025) 30–68. https://doi.org/10.1108/JKM-01-2025-0006.

[26] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, E. Vayena, AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, Minds Mach. (Dordr). 28 (2018) 689–707. https://doi.org/10.1007/s11023-018-9482-5.

[27] D.B. Resnik, M. Hosseini, The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool, AI and Ethics 5 (2025) 1499–1521. https://doi.org/10.1007/s43681-024-00493-8.

[28] J. Geerts, M. Nuyts, Responsible AI for scientific research, Pattern Recognit. Lett. 178 (2026). https://doi.org/10.1016/j.patrec.2025.01.004.

[29] Muneera Bano, Didar Zowghi, Pip Shea, Georgina Ibarra, Investigating Responsible AI for Scientific Research: An Empirical Stud, (2023). https://doi.org/https://doi.org/10.48550/arXiv.2312.09561.

[30] L. Ge, R. Agrawal, M. Singer, P. Kannapiran, J.A. De Castro Molina, K.L. Teow, C.W. Yap, J.A. Abisheganaden, Leveraging artificial intelligence to enhance systematic reviews in health research: advanced tools and challenges, Syst. Rev. 13 (2024). https://doi.org/10.1186/s13643-024-02682-2.

[31] S. Krakowski, Human-AI agency in the age of generative AI, Information and Organization 35 (2025). https://doi.org/10.1016/j.infoandorg.2025.100560.

[32] Fengchun. Miao, Stefania. Giannini, Wayne. Holmes, Guidance for generative AI in education and research, UNESCO, Paris, France, 2023.

[33] S. Akter, G. McCarthy, S. Sajib, K. Michael, Y.K. Dwivedi, J. D'Ambra, K.N. Shen, Algorithmic bias in data-driven innovation in the age of AI, Int. J. Inf. Manage. 60 (2021). https://doi.org/10.1016/j.ijinfomgt.2021.102387.

[34] A.H. Khan, D.K.J.B. Saini, T.H. Khan, B.K. Rai, A. Pimpalkar, G. Kumar, A quantum-driven multi-stage framework integrating variational entanglement, reinforcement learning, and federated explainability for climate-resilient farming, Sci. Rep. 15 (2025). https://doi.org/10.1038/s41598-025-22224-7.

[35] D. Kumar, A. Gupta, V.K. Gupta, A. Gupta, Aspect-Based Sentiment Analysis using Machine Learning and Deep Learning Approaches, International Journal on Recent and Innovation Trends in Computing and Communication 11 (2023) 118–138. https://doi.org/10.17762/ijritcc.v11i5s.6636.

[36] M.P. Geetha, D. Karthika Renuka, Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model, International Journal of Intelligent Networks 2 (2021) 64–69. https://doi.org/10.1016/j.ijin.2021.06.005.

[37] A. Jacovi, S. Swayamdipta, S. Ravfogel, Y. Elazar, Y. Choi, Y. Goldberg, Contrastive Explanations for Model Interpretability, (2021). http://arxiv.org/abs/2103.01378.

[38] L.H. Nazer, R. Zatarah, S. Waldrip, J.X.C. Ke, M. Moukheiber, A.K. Khanna, R.S. Hicklen, L. Moukheiber, D. Moukheiber, H. Ma, P. Mathur, Bias in artificial intelligence algorithms and recommendations for mitigation, PLOS Digital Health 2 (2023). https://doi.org/10.1371/journal.pdig.0000278.

[39] T.-Ștefan Rotaru, C. Amariei, Ethical issues in research with artificial intelligence systems, in: Ethics - Scientific Research, Ethical Issues, Artificial Intelligence and Education [Working Title], IntechOpen, 2023. https://doi.org/10.5772/intechopen.1001451.

[40] T. Hauer, Importance and limitations of AI ethics in contemporary society, Humanit. Soc. Sci. Commun. 9 (2022). https://doi.org/10.1057/s41599-022-01300-7.

[41] B. Mittelstadt, Principles alone cannot guarantee ethical AI, Nat. Mach. Intell. (2019) 501–507. https://doi.org/https://doi.org/10.1038/s42256-019-0114-4.
[42] R.J. Hanisch, I.S. Gilmore, A.L. Plant, Improving reproducibility in research: The role of measurement science, J. Res. Natl. Inst. Stand. Technol. 124 (2019). https://doi.org/10.6028/jres.124.024.
[43] H. Benbya, Artificial Intelligence in Organizations: Current State and Future Opportunities, n.d. https://aisel.aisnet.org/misqe/vol19/iss4/4.
[44] A. Rai, S. Sarker, EDITOR'S COMMENTS Next-Generation Digital Platforms: Toward Human-AI Hybrids, 2019. https://android-developers.googleblog.com/2018/01/how-we-fought-bad-apps-and-malicious.html;
[45] Z.C. Lipton, The Mythos of Model Interpretability, (2017). http://arxiv.org/abs/1606.03490.
[46] E.M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the dangers of stochastic parrots: Can language models be too big?, in: FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, Inc, 2021: pp. 610–623. https://doi.org/https://doi.org/10.1145/3442188.3445922.
[47] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, n.d. http://science.sciencemag.org/.
[48] A. Arora, M. Barrett, E. Lee, E. Oborn, K. Prince, Risk and the future of AI: Algorithmic bias, data colonialism, and marginalization, Information and Organization 33 (2023) 100478. https://doi.org/https://doi.org/10.1016/j.infoandorg.2023.100478.
[49] P.G. Hensel, Reproducibility and replicability crisis: How management compares to psychology and economics – A systematic review of literature, European Management Journal 39 (2021) 577–594. https://doi.org/10.1016/j.emj.2021.01.002.
[50] T. Susnjak, P. Hwang, N.H. Reyes, A.L.C. Barczak, T.R. McIntosh, S. Ranathunga, Automating Research Synthesis with Domain-Specific Large Language Model Fine-Tuning, (2024). https://doi.org/10.1145/3715964.
[51] C. Zednik, H. Boelsen, Scientific Exploration and Explainable Artificial Intelligence, Minds Mach. (Dordr). 32 (2022) 219–239. https://doi.org/10.1007/s11023-021-09583-6.
[52] A. Desai, M. Abdelhamid, N.R. Padalkar, What is Reproducibility in Artificial Intelligence and Machine Learning Research?, (2025). http://arxiv.org/abs/2407.10239.
[53] L. Nannini, M. Marchiori Manerba, I. Beretta, Mapping the landscape of ethical considerations in explainable AI research, Ethics Inf. Technol. 26 (2024). https://doi.org/10.1007/s10676-024-09773-7.
[54] S. Akter, G. McCarthy, Algorithmic bias in data-driven innovation, Int. J. Inf. Manage. 60 (2021). https://doi.org/10.1016/j.ijinfomgt.2021.102387.
[55] B. Shneiderman, Human-Centered Artificial Intelligence, Oxford University Press, 2022.
[56] B. Shneiderman, Human-Centered Artificial Intelligence: Three Fresh Ideas, AIS Transactions on Human-Computer Interaction 12 (2020) 109–124. https://doi.org/10.17705/1thci.00131.
[57] N. Rafaqat Ahmad, LEX LOCALIS-JOURNAL OF LOCAL SELF-GOVERNMENT HUMAN-AI COLLABORATION IN KNOWLEDGE WORK: PRODUCTIVITY, ERRORS, AND ETHICAL RISK, n.d.
[58] K. Sowa, A. Przegalinska, From Expert Systems to Generative Artificial Experts: A New Concept for Human-AI Collaboration in Knowledge Work, 2025.
[59] J. Holmström, From AI to digital transformation: The AI readiness framework, Bus. Horiz. 65 (2022) 329–339. https://doi.org/10.1016/j.bushor.2021.03.006.
[60] J. de la Torre-López, A. Ramírez, J.R. Romero, Artificial intelligence to automate the systematic review of scientific literature, Computing 105 (2023) 2171–2194. https://doi.org/10.1007/s00607-023-01181-x.
[61] B. Shneiderman, Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy, 2020. https://arxiv.org/abs/2002.04087v1.