ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor = 8.35, January-June 2026, Published on 28th February 2026

Artificial Intelligence in Research: Opportunities, Challenges, and Ethical Considerations

Mr. Rajdeep Singh, Research Scholar, Department of Political Science, Chaudhary Devi Lal University, Sirsa
Email: rajdeep151507@gmail.com

Abstract

Artificial Intelligence (AI) is rapidly transforming the landscape of academic and scientific research by enhancing efficiency, accuracy, and innovation across disciplines. From automated literature reviews and data mining to predictive modelling and advanced analytics, AI-driven tools are reshaping research methodologies and accelerating knowledge production. These technologies enable researchers to handle large datasets, identify complex patterns, and generate insights that were previously difficult to obtain. However, the integration of AI in research also presents significant challenges, including issues of algorithmic bias, data privacy, transparency, intellectual property, and academic integrity. Ethical concerns surrounding authorship, accountability, and over-reliance on automated systems require careful governance and regulatory frameworks. This paper critically examines the opportunities AI offers for improving research productivity and interdisciplinary collaboration while highlighting the limitations and ethical dilemmas that accompany its adoption. It argues for a balanced, responsible, and policy-guided approach to ensure that AI strengthens, rather than compromises, the integrity and credibility of research.

Keywords: Artificial Intelligence, research methodology, algorithmic bias, academic integrity, research ethics, data privacy, machine learning

Introduction

The emergence and rapid advancement of Artificial Intelligence (AI) represents one of the most profound technological shifts in the history of human knowledge production. Over the past decade, AI has migrated from the domain of computer science laboratories into virtually every branch of academic inquiry, from the natural sciences and medicine to the social sciences, humanities, and policy research. This migration has been driven by exponential growth in computational power, the proliferation of large digitised datasets, and the development of sophisticated machine learning (ML) algorithms capable of performing tasks once thought to require human intelligence (Russell & Norvig, 2020). The implications for research are immense, touching on how scholars discover literature, design studies, collect and analyse data, generate hypotheses, and even communicate their findings.

In academic settings worldwide, AI tools are being deployed at every stage of the research pipeline. Natural Language Processing (NLP) systems can now scan thousands of academic articles within minutes to synthesise existing evidence on a given topic (Tsafnat et al., 2014). Deep learning models can detect patterns in genomic data, medical images, climate models, and social network structures with a level of precision and scale that no human researcher could achieve unaided (LeCun et al., 2015). Platforms powered by AI are reshaping peer review, grant assessment, and citation analysis, fundamentally altering the incentive structures and gatekeeping mechanisms that have long governed scholarship.

Yet this technological revolution is not without its shadows. As AI becomes more deeply embedded in research processes, scholars and policymakers are increasingly alert to a range of serious challenges. Algorithmic systems can encode and amplify existing biases present in training data, producing distorted findings with wide-reaching consequences (Barocas et al., 2019). The opacity of many AI models—their so-called 'black box' character—raises fundamental questions about the reproducibility and verifiability of AI-assisted research (Rudin, 2019). The growing use of AI writing tools and autonomous systems to generate research content poses direct challenges to established norms of authorship, intellectual property, and academic honesty (Stokel-Walker & Van Noorden, 2023). Meanwhile, the vast quantities of personal and sensitive data required to train many AI systems create substantial risks of privacy violation and data misuse (Mittelstadt & Floridi, 2016).

This paper provides a comprehensive critical examination of the opportunities and challenges presented by the integration of AI into research. Drawing on a wide body of scholarly literature, it explores the transformative potential of AI across different research functions, identifies the most pressing ethical and practical concerns, and reflects on the governance frameworks needed to ensure responsible adoption. The paper argues that maximising the benefits of AI in research requires more than technological optimism; it demands careful institutional design, transparent practice, and a renewed commitment to the foundational values of scientific integrity and social responsibility.

Artificial Intelligence in Research: An Overview

AI is not a single technology but a constellation of computational methods that simulate aspects of human cognition. For research purposes, the most relevant capabilities include machine learning (ML), deep learning, natural language processing, computer vision, and autonomous systems (Goodfellow et al., 2016). Each of these has found important applications in academic and scientific inquiry.

Machine learning, the branch of AI concerned with enabling computers to learn from data without explicit programming, has become foundational to modern data-driven research. Supervised learning techniques allow researchers to build predictive models from labelled training data, while unsupervised learning methods—such as clustering and dimensionality reduction—enable the discovery of hidden structures within complex datasets (Bishop, 2006). These capabilities have transformed disciplines including epidemiology, economics, political science, and psychology, where large observational datasets can now be analysed with unprecedented depth and speed.

Natural Language Processing has opened up entirely new research possibilities in text-heavy fields. Systematic reviews and meta-analyses, which traditionally required teams of researchers months of work to locate, screen, and synthesise relevant literature, can now be substantially automated using NLP-powered tools (Marshall & Wallace, 2019). Sentiment analysis, topic modelling, and semantic similarity algorithms allow researchers to extract meaning from vast textual corpora—whether historical archives, social media data, legal documents, or published articles—at a scale impossible through manual effort alone (Blei et al., 2003).

Computer vision technologies have revolutionised research in medicine, biology, astronomy, and environmental science. Convolutional neural networks can identify cancerous lesions in radiological images with diagnostic accuracy rivalling that of experienced clinicians (Esteva et al., 2017). In ecology, AI-powered image recognition systems can count and classify species in wildlife camera trap photographs, enabling real-time monitoring of biodiversity at continental scales. In astronomy, deep learning models analyse telescope data to identify exoplanets, gravitational lenses, and transient phenomena that would otherwise go undetected in the torrent of observational data.

Generative AI systems—including large language models (LLMs) such as GPT-4 and similar architectures—represent the most recent and perhaps most disruptive wave of AI development for researchers. These systems can draft literature reviews, generate hypotheses, write and debug code, translate scientific documents, and produce summaries of complex research findings (Brown et al., 2020). While these capabilities offer substantial productivity gains, they also introduce novel risks around intellectual authorship and the authenticity of knowledge production, which are examined in detail later in this paper.

Opportunities: How AI Artificial Intelligence (AI) is Transforming Research by various means

Accelerating Literature Review and Knowledge Synthesis

One of the most tangible benefits of AI in research is its ability to dramatically accelerate the process of literature review and systematic knowledge synthesis. The volume of published scientific literature has grown exponentially in recent decades, with an estimated five million peer-reviewed articles published annually across all disciplines (Johnson et al., 2018). For individual researchers, keeping pace with this output within even a narrowly defined subfield has become practically impossible. AI-powered literature review tools such as Elicit, Semantic Scholar, and Rayyan leverage NLP algorithms to assist researchers in identifying relevant studies, extracting key findings, and mapping conceptual relationships across thousands of documents simultaneously.

Tsafnat et al. (2014) demonstrated that machine learning-based screening tools could reduce the human effort required in systematic review by up to 70% while maintaining acceptable levels of sensitivity. Subsequent studies have consistently supported this finding across different medical and social science domains. Beyond efficiency gains, AI-assisted literature review can improve comprehensiveness by reducing the risk that important studies are missed due to human fatigue or incomplete search strategies. The development of automated evidence synthesis tools promises to accelerate the translation of research findings into clinical guidelines, policy recommendations, and industry practice.

Enhanced Data Analysis and Pattern Recognition

Modern research increasingly depends on the analysis of large, high-dimensional datasets that exceed the analytical capacity of traditional statistical methods. AI's ability to identify complex, non-linear patterns within such data represents a fundamental expansion of researchers' analytical toolkit. In genomics, deep learning models trained on sequencing data have identified genetic variants associated with complex diseases that remained invisible to conventional genome-wide association studies (Ching et al., 2018). In social science, ML algorithms applied to administrative data have revealed heterogeneous treatment effects and subgroup dynamics that aggregate analyses would mask.

Beyond pattern recognition, AI enables causal inference at new scales. Natural experiments involving millions of individuals—employment records, educational trajectories, health outcomes—can now be analysed using ML-augmented causal frameworks that account for confounding in ways traditional regression approaches cannot (Athey & Imbens, 2019). In climate science, neural network models trained on satellite imagery and atmospheric sensor data provide more accurate regional climate projections than previous generation models, informing both adaptation planning and mitigation policy (Reichstein et al., 2019).

Interdisciplinary Collaboration and Open Science

AI is also acting as a catalyst for interdisciplinary research by providing shared methodological tools and infrastructures that span traditional disciplinary boundaries. Computational methods originally developed in computer science—including graph neural networks, transformer architectures, and reinforcement learning—are now routinely deployed in biology, linguistics, economics, and the humanities. This methodological convergence is creating new hybrid disciplines such as computational social science, digital humanities, and AI-assisted drug discovery.

The open science movement, which advocates for the public sharing of research data, code, and findings, has found a powerful ally in AI. AI-powered platforms can facilitate the standardisation and interoperability of research datasets, making it easier for researchers across institutions and countries to collaborate and replicate each other's work (Wilkinson et al., 2016). Tools such as automated data quality checks and metadata generation reduce the technical barriers to data sharing, while AI-powered translation systems are beginning to erode the linguistic barriers that have historically confined much global research output to English-language publication venues.

Drug Discovery and Biomedical Innovation

Perhaps nowhere is the transformative potential of AI in research more vividly illustrated than in drug discovery and biomedical science. The traditional drug development pipeline—from

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor = 8.35, January-June 2026, Published on 28th February 2026

target identification to clinical approval—typically spans fifteen years and costs over one billion dollars, with a high rate of failure at each stage (DiMasi et al., 2016). AI is beginning to compress this timeline dramatically. AlphaFold, developed by DeepMind, solved the fifty-year-old problem of protein structure prediction with near-atomic accuracy, providing free access to structural predictions for virtually every known protein (Jumper et al., 2021). This achievement has profound implications for understanding disease mechanisms and identifying novel therapeutic targets.

Machine learning models trained on existing pharmacological data can screen billions of candidate compounds for bioactivity, toxicity, and drug-like properties in silico before any experimental synthesis is required (Stokes et al., 2020). During the COVID-19 pandemic, AI tools contributed to the rapid identification of potential antiviral candidates and assisted in the analysis of viral evolutionary trajectories, demonstrating the real-world value of AI-accelerated research in conditions of public health emergency.

Challenges and Limitations: There are several challenges encountered while AI used in the academic research

Algorithmic Bias and Fairness

Despite its transformative potential, the deployment of AI in research is accompanied by a range of serious challenges that demand rigorous attention. Chief among these is the problem of algorithmic bias. AI systems learn from historical data and, in doing so, can inherit, encode, and amplify the biases, inequities, and blind spots present in that data (Barocas et al., 2019). In medical research, datasets used to train diagnostic AI models have frequently overrepresented certain demographic groups—most notably white male patients in high-income countries—resulting in models that perform less accurately for women, racial minorities, and patients in low- and middle-income settings (Obermeyer et al., 2019).

The consequences of biased AI in research extend beyond individual studies. When biased models are used as foundations for further research, the distortions compound over time, producing bodies of literature that systematically misrepresent the populations and phenomena they purport to describe. In the social sciences, predictive policing algorithms trained on historically biased crime data have been critiqued for reinforcing racial disparities in law enforcement (Angwin et al., 2016). Addressing algorithmic bias requires not only technical interventions—such as debiasing algorithms and diverse training data—but also structural commitments to inclusive data collection practices and meaningful community involvement in the design of AI research tools.

Transparency and Reproducibility

A core principle of scientific research is that findings should be transparent, reproducible, and open to scrutiny. This principle is significantly challenged by the opacity of many contemporary AI systems. Deep learning models in particular are often characterised as 'black boxes': their internal representations and decision-making processes are not readily interpretable by human researchers, making it difficult to explain why a given model produces a given output (Rudin, 2019). This opacity poses fundamental problems for research credibility. If a model's reasoning cannot be examined and understood, its conclusions cannot be properly evaluated, and errors or biases within the model may go undetected.

Reproducibility is also a major concern. Many AI-based research findings depend critically on specific hyperparameter choices, random seeds, and computational environments that are not fully documented in published papers. Several high-profile studies in machine learning and AI research have been found to be difficult or impossible to reproduce, raising concerns about the robustness of findings reported in this literature (Hutson, 2018). The AI research community has begun to respond to these concerns through the development of model cards, datasheets for datasets, and reporting standards that improve documentation and transparency (Mitchell et al., 2019; Gebru et al., 2021).

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor = 8.35, January-June 2026, Published on 28th February 2026

Data Privacy and Security

Many of the most powerful AI research applications depend on access to large quantities of sensitive personal data. Medical AI systems require patient health records; social science AI applications draw on social media data, administrative records, and survey responses; educational AI tools collect detailed behavioural and performance data on students. The aggregation and use of such data for research purposes raises profound questions about privacy, consent, and the appropriate limits of data use (Mittelstadt & Floridi, 2016).

Even when researchers obtain appropriate ethical approval for data use, the risk of re-identification—the process by which supposedly anonymised data can be linked to specific individuals using additional information—presents ongoing challenges. Studies have demonstrated that supposedly anonymised medical records can be re-identified with high accuracy using auxiliary data sources such as public social media profiles (Sweeney, 2002). As AI systems become increasingly adept at integrating and analysing disparate data sources, the technical barriers to re-identification continue to fall, creating growing tensions between the value of data-driven research and the fundamental right to privacy.

Over-Reliance and Deskilling

A more subtle but potentially serious challenge associated with the integration of AI in research is the risk of over-reliance and associated deskilling. As AI systems increasingly automate cognitive tasks previously performed by human researchers—such as literature search, data coding, statistical analysis, and even writing—there is a risk that researchers may gradually lose the skills and intuitions necessary to critically evaluate AI outputs, identify errors, and exercise independent scholarly judgment (Sweeney & Criado-Perez, 2021). This concern is particularly acute for early-career researchers who may develop their scholarly practices in environments where AI assistance is ubiquitous.

Over-reliance on AI-generated literature reviews, for example, risks producing researchers who are less familiar with primary sources and less capable of the deep disciplinary engagement necessary for genuine theoretical innovation. Automated data analysis tools may reduce researchers' understanding of the statistical assumptions underlying their models, increasing the risk of methodological errors that go undetected. Addressing this challenge requires pedagogical innovations that ensure researchers develop robust foundational competencies even as they learn to leverage AI tools effectively.

Ethical Considerations: Use for AI in research Raise several ethical concerns.

Authorship and Academic Integrity

The proliferation of generative AI tools capable of producing fluent, scholarly-sounding text has precipitated a crisis in academic authorship norms. Large language models such as GPT-4 can generate literature reviews, discussion sections, and even full research papers on demand, raising acute questions about what it means to be the author of a piece of research and what responsibilities authorship confers (Stokel-Walker & Van Noorden, 2023). Several high-profile journals and conferences have received and even provisionally accepted submissions that were wholly or substantially generated by AI, prompting urgent policy responses from publishers, universities, and scholarly societies.

The responses have varied considerably. Some journals have prohibited the use of AI language models as co-authors on the grounds that authorship requires accountability that AI systems cannot provide (Thorp, 2023). Others have adopted more permissive stances, requiring only disclosure of AI tool use rather than prohibiting it outright. The absence of a universal standard creates risks of inconsistency and arbitrage, where researchers exploit the most permissive policies to engage in practices that undermine the integrity of scholarship. Beyond the question of text generation, AI tools are increasingly being used to generate synthetic data, manipulate research images, and automate the production of statistical analyses, all of which raise serious concerns about data fabrication and research fraud.

Intellectual Property and Copyright

The use of AI in research also creates complex and unresolved questions of intellectual property and copyright. Many AI systems, including large language models and image generation systems, are trained on vast datasets of copyrighted text, images, and other material, often without the explicit consent of the copyright holders (Sag, 2018). Researchers who use such systems to produce content for publication may inadvertently reproduce copyrighted material or find themselves entangled in legal disputes about the ownership of AI-generated outputs.

The question of who owns the intellectual property in research outputs that are substantially produced with AI assistance is particularly contentious. Traditional intellectual property frameworks assign ownership to human creators; AI systems have no legal personhood and cannot hold intellectual property rights. However, the human contribution to an AI-generated work may be minimal—amounting to little more than the provision of a prompt—calling into question whether the resulting output should attract copyright protection at all (Grimmelmann, 2016). These questions have significant implications for the incentive structures that underpin academic research and innovation.

Accountability and Governance

Effective governance of AI in research requires clear frameworks for accountability—mechanisms that identify who is responsible when AI systems produce harmful, misleading, or erroneous research outputs. The distributed character of AI research systems complicates accountability: the developers of an AI tool, the researchers who deploy it, the institutions that fund its use, and the journals that publish its outputs may all bear some responsibility for problematic outcomes, but the apportionment of that responsibility is rarely clear (Dafoe, 2018).

Several international bodies have developed principles and guidelines for the ethical governance of AI, including the OECD Principles on Artificial Intelligence (OECD, 2019) and UNESCO's Recommendation on the Ethics of AI (UNESCO, 2021). These frameworks emphasise values including transparency, accountability, human oversight, non-discrimination, and respect for human rights. However, the translation of these high-level principles into enforceable norms for specific research contexts remains incomplete. Individual institutions have developed their own policies on AI use in research, but these vary widely and often lag behind the pace of technological development.

Environmental and Resource Justice

An often-overlooked ethical dimension of AI in research is its environmental impact and implications for resource justice. Training large-scale AI models requires substantial computational resources and correspondingly large amounts of energy. Strubell et al. (2019) estimated that training a single large NLP model can generate carbon emissions comparable to the entire lifetime carbon footprint of five average American automobiles. As AI systems grow larger and more complex, their energy demands are increasing, raising questions about the environmental sustainability of AI-intensive research practices.

Furthermore, access to the computational infrastructure required for state-of-the-art AI research is highly unequally distributed, both within and between countries. A small number of wealthy institutions in high-income countries—primarily large technology companies and elite universities—control the most powerful AI systems and the largest proprietary datasets. This concentration of AI research capacity risks exacerbating existing inequalities in global knowledge production and reinforcing the epistemic dominance of already-privileged institutions and societies (Mohamed et al., 2020).

Towards Responsible Integration: Policy and Practice

Given the magnitude of both the opportunities and the challenges associated with AI in research, the development of robust and adaptive governance frameworks is an urgent priority. Effective governance must operate at multiple levels—individual researcher, institutional, disciplinary, national, and international—and must be sufficiently flexible to keep pace with rapidly evolving technological capabilities.

At the level of research practice, several principles can guide responsible AI integration. Transparency requires that researchers fully disclose when and how AI tools have been used in the production of research, including in literature searches, data analysis, and writing. Validation requires that AI-generated findings be subject to rigorous human scrutiny and, where possible, independent replication before being treated as reliable (National Academies of Sciences, Engineering, and Medicine, 2019). Inclusivity requires that AI tools be evaluated for bias and differential performance across different populations before being deployed in consequential research contexts. Sustainability requires that researchers consider the environmental footprint of their AI use and seek to minimise unnecessary computational expenditure.

Institutional responsibilities include the provision of training and education to ensure that researchers at all career stages develop the competencies needed to use AI tools critically and responsibly. Universities and research funders should invest in AI literacy programmes that address both technical skills and ethical understanding, and should ensure that institutional policies on AI use are clearly communicated and consistently enforced. Research funding bodies should require grant applicants to address AI use in their data management and research ethics plans, and should support the development of shared standards and best practices within and across disciplines (Jobin et al., 2019).

At the disciplinary level, scholarly societies and journal editors play a critical role in setting norms around AI use in research. The development of clear, consistent, and evidence-based authorship policies is particularly urgent. Peer review processes should be updated to include assessment of AI-related methodological transparency, and editors should be equipped to identify and address potential cases of AI-assisted misconduct. The establishment of shared repositories for AI research tools and associated datasets would facilitate replication and independent verification of AI-assisted findings.

Internationally, the development of binding norms and regulatory frameworks for AI in research will require sustained multilateral cooperation. While achieving global consensus on specific rules may be challenging given the diversity of national legal, cultural, and institutional contexts, agreement on core principles—transparency, accountability, non-discrimination, and human oversight—provides a foundation for more detailed normative development (Calo, 2017). Regulatory initiatives such as the European Union's Artificial Intelligence Act, which establishes a risk-based framework for the governance of high-risk AI applications, offer instructive models for balancing innovation with protection of fundamental rights and values (European Commission, 2021).

Conclusion

Artificial Intelligence represents both a remarkable opportunity and a profound challenge for academic and scientific research. Its capacity to accelerate knowledge discovery, enhance analytical power, foster interdisciplinary collaboration, and address complex global problems is genuinely transformative. At the same time, the integration of AI into research processes raises serious questions about bias, transparency, privacy, intellectual property, academic integrity, accountability, and resource justice that cannot be resolved by technological innovation alone. The argument advanced in this paper is that realising the benefits of AI in research while managing its risks requires a commitment to responsible integration rather than either uncritical adoption or reflexive resistance. Researchers, institutions, funders, publishers, and policymakers all have vital roles to play in developing the norms, frameworks, and practices needed to ensure that AI strengthens rather than undermines the integrity and credibility of the research enterprise. The stakes are high. Research is one of humanity's most important collective activities—the primary means by which knowledge is produced, tested, and refined in the service of human flourishing. The values that have historically underpinned research—honesty, rigour, transparency, accountability, and respect for persons—are not outdated remnants of a pre-digital era. They are enduring foundations that must be actively protected and reaffirmed as the technological landscape of research is transformed. Getting the governance of AI in research right is not merely a technical or regulatory challenge; it is an ethical and civilisational imperative.

References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. Annual Review of Economics, 11(1), 685–725. https://doi.org/10.1146/annurev-economics-080217-053433

Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. fairmlbook.org.

Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Calo, R. (2017). Artificial intelligence policy: A primer and roadmap. UC Davis Law Review, 51(2), 399–435.

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., Ferrero, E., Agapow, P.-M., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G. L., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., . . . Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of the Royal Society Interface, 15(141), Article 20170387. https://doi.org/10.1098/rsif.2017.0387

Dafoe, A. (2018). AI governance: A research agenda. Future of Humanity Institute, University of Oxford.

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 47, 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

European Commission. (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). European Commission. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Grimmelmann, J. (2016). Copyright for literate robots. Iowa Law Review, 101(2), 657–681.

Hutson, M. (2018). Artificial intelligence faces reproducibility crisis. Science, 359(6377), 725–726. https://doi.org/10.1126/science.359.6377.725

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2
Johnson, R., Watkinson, A., & Mabe, M. (2018). The STM report: An overview of scientific and scholarly publishing (5th ed.). International Association of Scientific, Technical and Medical Publishers.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539
Marshall, I. J., & Wallace, B. C. (2019). Toward systematic review automation: A practical guide to using machine learning tools in research synthesis. Systematic Reviews, 8(1), Article 163. https://doi.org/10.1186/s13643-019-1074-9
Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–229. https://doi.org/10.1145/3287560.3287596
Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. Science and Engineering Ethics, 22(2), 303–341. https://doi.org/10.1007/s11948-015-9652-2
Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy & Technology, 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8
National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and replicability in science. The National Academies Press. https://doi.org/10.17226/25303
Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453. https://doi.org/10.1126/science.aax2342
OECD. (2019). Recommendation of the Council on artificial intelligence. Organisation for Economic Co-operation and Development. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1
Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x
Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach (4th ed.). Pearson.
Sag, M. (2018). The new legal landscape for text mining and machine learning. Journal of the Copyright Society of the U.S.A., 66(2), 291–366.
Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. Nature, 614(7947), 214–216. https://doi.org/10.1038/d41586-023-00340-6
Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688–702. https://doi.org/10.1016/j.cell.2020.01.021
Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep

learning in NLP. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3645–3650. https://doi.org/10.18653/v1/P19-1355

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5), 557–570. https://doi.org/10.1142/S0218488502001648

Thorp, H. H. (2023). ChatGPT is fun, but not an author. Science, 379(6630), 313. https://doi.org/10.1126/science.adg7879

Tsafnat, G., Glasziou, P., Choong, M. K., Dunn, A., Galgani, F., & Coiera, E. (2014). Systematic review automation technologies. Systematic Reviews, 3(1), Article 74. https://doi.org/10.1186/2046-4053-3-74

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3(1), Article 160018. https://doi.org/10.1038/sdata.2016.18