

Research Article

Research Integrity in the Age of Artificial Intelligence: Authorship, Transparency, and AccountabilitySwati Joshi^{1,*}, Ravi Ranjan², Anirudh Gupta³ and Divya Nigam⁴¹Department of Computer Science, P. V. G's College of Science and Commerce, Pune, India²Department of Computer Science and Engineering, Government Engineering College Banka, India³Department of Applied Sciences, Suresh Gyan Vihar University, India⁴Department of Education, Bundelkhand University Jhansi, India

*Corresponding Author: Swati Joshi. Email: jsp15@rediffmail.com

Received: 29/01/2026; Revised: 16/02/2026; Accepted: 18/03/2026; Published: 30/04/2026.

DOI: <https://doi.org/10.69996/fmep.2026016>

Abstract: The rapid diffusion of generative artificial intelligence (AI) into scholarly work has unsettled long-standing norms of research integrity. Tools that can draft manuscripts, propose analyses, and generate plausible text and data at scale create value for researchers but also blur the lines of who authored a work, how it was produced, and who is answerable for it. This conceptual paper examines research integrity in the age of AI through three interlocking dimensions: authorship, transparency, and accountability. It argues that AI, lacking legal personhood and the capacity to take responsibility, cannot be an author; that any substantive use of AI must be disclosed if the reproducibility and trustworthiness of research are to be preserved; and that accountability for AI-assisted work must remain firmly with identifiable human agents. Building on emerging editorial and ethical guidance, the paper presents an integrated framework in which these three principles rest on a foundation of enabling mechanisms, namely disclosure norms, human oversight, institutional and editorial governance, AI literacy, and equitable access. It further maps the principal integrity challenges AI introduces to corresponding safeguards. The paper concludes that AI is best understood as an augmenting instrument whose responsible use depends less on the technology itself than on the human and institutional structures that surround it, and it calls for empirical study of how disclosure and accountability norms take hold in practice.

Keywords: Research Integrity, Artificial Intelligence, Generative AI, Authorship, Transparency, Accountability, Research Ethics, Scholarly Publishing.

1 Introduction

Research integrity, the commitment to honesty, rigour, transparency, and accountability in the conduct and reporting of research, is the foundation on which the trustworthiness of science rests. It is sustained not by any single rule but by a web of shared norms: that authors are those who made genuine intellectual contributions and can answer for the work; that methods are reported openly enough to be scrutinised and reproduced; and that responsibility for what is claimed lies with identifiable people. The rapid arrival of generative artificial intelligence (AI) has placed pressure on each of these norms at once.

Within a short period, AI systems capable of producing fluent scholarly text, proposing analyses, summarising literature, and generating plausible data have moved from novelty to everyday instrument. These capabilities offer real benefits, from language support that widens participation to efficiencies that free researchers for higher-order work. But they also unsettle settled assumptions. When a manuscript is drafted with the substantial help of a machine, who is its author? When AI shapes a study's methods or text, how much of that use must be revealed? And when an AI-assisted work turns out to contain fabricated citations or flawed reasoning, who is accountable?



This paper addresses these questions by examining research integrity in the age of AI through three interlocking dimensions: authorship, transparency, and accountability. It is a conceptual and integrative contribution rather than an empirical study; its aim is to synthesise current developments and emerging editorial guidance into a coherent framework for responsible practice. The central argument is that AI is best understood as an augmenting instrument, not an autonomous author or investigator, and that safeguarding integrity depends less on the technology than on the human judgement and institutional structures that surround its use. Section 2 sets out the conceptual terrain. Sections 3, 4, and 5 examine authorship, transparency, and accountability in turn. Section 6 presents an integrated framework and a mapping of challenges to safeguards. Section 7 discusses tensions and open problems, Section 8 offers recommendations, and Section 9 concludes.

2. Research Integrity in the AI Era: Conceptual Terrain

Traditional accounts of research integrity emphasise the avoidance of fabrication, falsification, and plagiarism, together with positive commitments to rigour, openness, and responsible conduct. These commitments assume a human researcher who intends, decides, and can be held to account. Generative AI complicates that assumption in a specific way: it can produce outputs that resemble the products of human intellectual labour, including prose, code, and apparent evidence, without any understanding, intention, or capacity for responsibility. The outputs are optimised for plausibility rather than truth, which means that fluent, confident text may be wholly or partly fabricated.

This creates a mismatch between the surface form of AI outputs and the normative structure of research. A citation generated by a model looks like a citation but may refer to nothing; an analysis suggested by a model looks authoritative but may be unfounded. The integrity of research therefore comes to depend, more than before, on the discipline with which humans verify, disclose, and take responsibility for what AI contributes. The three dimensions examined below can be read as the points at which this discipline is most consequential.

3. Authorship in the Age of AI

Authorship is both a credit and a responsibility. To be named as an author is to claim intellectual contribution and, crucially, to accept accountability for the work, including the ability to approve the final version and to respond to questions and challenges after publication. It is on this second point that the case against AI authorship is decisive. An AI system cannot take responsibility for a work, cannot approve it in any meaningful sense, and cannot answer for it; it has neither the legal standing nor the moral agency that authorship presupposes. Editorial and ethics bodies have converged on this conclusion, holding that AI tools may not be listed as authors.

This does not mean AI has no legitimate place in the production of scholarship. The emerging distinction is between using AI as an aid, for example to improve language, structure, or clarity, and delegating to it the intellectual substance of the work. The former is broadly accepted provided it is disclosed; the latter erodes the very contribution that authorship is meant to recognise. A related concern is the integrity of the human authorship that remains: over-reliance on generative tools can hollow out the skills and judgement that scholarly work is meant to embody, and undisclosed machine drafting is a form of misrepresentation. Preserving authorship in the age of AI therefore requires both a clear rule, that AI is a tool and not an author, and a culture in which humans retain genuine intellectual ownership of what they publish.

4. Transparency and Disclosure

If authorship answers the question of who is credited, transparency answers the question of how a work was produced. Reproducibility and scrutiny, the mechanisms by which science corrects itself, depend on a sufficiently open account of methods. Undisclosed AI use undermines this in two ways. First, it hides a step in the research process that may have shaped the results, making the work harder to evaluate and to reproduce. Second, because AI systems are frequently updated, often proprietary, and sometimes non-deterministic, an analysis produced with a given tool at a given time may not be reproducible later unless the tool, version, and manner of use are recorded.

The response that has emerged across publishers, funders, and editorial bodies is a norm of disclosure: researchers should state where and how AI was used, in both the conduct of the research and the preparation of the manuscript. Good disclosure identifies the tools and versions, the tasks for which they were used, and, where relevant, the prompts or configurations involved, so that a reader can weigh the machine's contribution. Disclosure is not an admission of wrongdoing but a precondition of trust; it allows AI to be used openly rather than covertly, and it keeps the record honest. The practical challenge is to make disclosure specific enough to be meaningful without becoming so burdensome that it is ignored, a balance that journal policies are still calibrating.

5. Accountability and Responsibility

Accountability is the dimension that binds the others together. Even with clear authorship rules and honest disclosure, the question remains of who answers when AI-assisted research goes wrong, when a fabricated citation is discovered, a flawed analysis is published, or a claim proves unfounded. The tempting move is to displace responsibility onto the tool, a rhetorical "the AI did it" that mirrors what has been called the responsibility gap: the worry that as decisions are delegated to autonomous systems, no human remains clearly answerable for the outcome.

In research, this gap must be firmly closed. Because AI cannot bear responsibility, the human agents who chose to use it, who accepted its outputs, and who submitted the work retain full accountability for it. This principle has a clarifying effect on practice: if researchers know they remain answerable for everything AI contributes, they have a strong incentive to verify its outputs, to disclose its use, and to treat it as an instrument under their control rather than a source of authority in its own right. Institutions and publishers reinforce this by directing accountability to named individuals and by declining to accept the involvement of a tool as an excuse. Accountability, in short, is what makes the augmentation model workable: AI may do much of the labour, but people remain responsible for the result.

6. An Integrated Framework

The three dimensions examined above are not independent; they reinforce one another and share a common foundation. Authorship without accountability is empty, accountability without transparency is unenforceable, and transparency without a culture that values it is merely nominal. Figure 1 presents these relationships as an integrated framework in which authorship, transparency, and accountability operate as pillars upholding research integrity, and in which those pillars themselves rest on a foundation of enabling mechanisms.

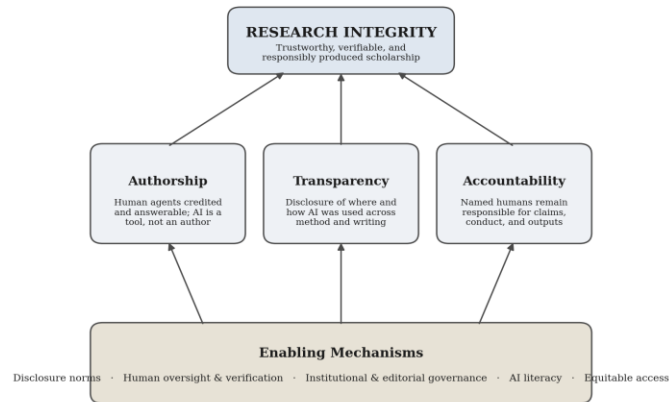


Figure 1: An integrated framework for research integrity in the age of AI. Three pillars (authorship, transparency, accountability) uphold research integrity and are themselves supported by a foundation of enabling mechanisms.

The foundation layer is decisive. Disclosure norms make transparency operational; human oversight and verification convert accountability from a slogan into a practice; institutional and editorial governance supply the rules and consequences that give the principles force; AI literacy ensures that researchers understand the capabilities and limits of the tools they use; and equitable access guards against the framework benefiting only the well-resourced. None of these mechanisms is sufficient alone, and none is supplied by the technology itself; each must be built into research practice and policy.

Table 1 complements the framework by mapping the principal integrity challenges that AI introduces to the safeguards that answer them. Read together, the figure and the table express the paper’s central claim: the risks AI poses to research integrity are real but tractable, and each has a corresponding human and institutional response.

Table 1: Dimensions of research-integrity challenges posed by AI and corresponding safeguards

Dimension	Principal challenge introduced by AI	Safeguard / responsible response
Authorship	Attribution of AI’s contribution; risk of listing AI as an author or of hidden machine-generated content	AI cannot be an author; credit and answerability rest with humans; substantive AI use is acknowledged
Transparency	Undisclosed AI use obscures how a work was produced and undermines reproducibility	Disclose tools, versions, and the nature of AI use across both method and writing
Accountability	Diffusion of responsibility; the “the AI did it” defence and the responsibility gap	Named human agents remain responsible for all claims, conduct, and outputs
Data & analysis integrity	Fabrication and “hallucination” of citations, statistics, and results	Independent verification of AI outputs against primary sources before use
Equity	Unequal access to capable tools risks a new divide among researchers	Affordable or open access, capacity building, and attention to global inclusion

7. Tensions and Open Problems

The framework does not dissolve every difficulty. Several tensions remain live. One concerns the boundary of disclosure: if trivial uses of AI, such as routine grammar checking, need not be reported while substantive uses must be, the line between the two is genuinely unclear and is drawn differently across journals. A second tension concerns detection. As AI-generated text becomes harder to distinguish from human writing, enforcement of disclosure norms cannot rely on detection alone and must instead depend on a culture of honest reporting, supported but not replaced by technical tools.

A third tension is equity. Measures intended to protect integrity, such as restrictions on AI use or demands for elaborate disclosure, may fall unevenly, disadvantaging researchers who rely on AI for legitimate language support. A fourth concerns pace: editorial policies and institutional codes evolve more slowly than the tools they govern, so that guidance is often reactive. These tensions do not undermine the framework, but they mark the places where judgement, rather than rule-following, will continue to be required, and where empirical study of actual practice would be especially valuable.

8. Recommendations for Responsible Practice

Several practical commitments follow from the analysis. Researchers should disclose substantive AI use in a specific and honest manner, verify all AI-generated content against primary sources before relying on it, and treat AI as an instrument for which they remain accountable rather than as an author or authority. Institutions and publishers should maintain clear, current policies on acceptable use and disclosure, protect the confidentiality of unpublished work, and direct accountability unambiguously to named individuals. The wider community should invest in AI literacy, so that researchers can use these tools competently and critically, and in equitable access, so that the benefits and the obligations of AI use are shared. Finally, integrity infrastructure, including provenance standards and disclosure mechanisms, should be strengthened to protect the scholarly record on which both human researchers and AI systems increasingly depend.

9. Conclusion

Artificial intelligence has become part of the machinery of research, and in doing so has tested the norms that make research trustworthy. This paper has argued that the challenge is best met not by resisting AI nor by surrendering to it, but by holding firm to three principles: that authorship, with its inseparable burden of accountability, belongs to humans; that the use of AI must be transparent if research is to remain reproducible and open to scrutiny; and that responsibility for AI-assisted work rests with identifiable people. Framed as an integrated structure resting on enabling mechanisms, these principles offer a workable path for preserving integrity while realising the benefits AI can bring.

The decisive variable throughout is human and institutional rather than technical. AI can draft, suggest, and accelerate, but it cannot be honest, take responsibility, or care about the truth; those remain the province of researchers and the communities that hold them to account. As a conceptual contribution, this paper advances propositions that invite empirical examination: how disclosure norms are actually adopted, how accountability is assigned in practice when AI-assisted work fails, and how equitably the benefits and burdens of these tools are distributed. Pursuing those questions is the next step in ensuring that research integrity endures in the age of artificial intelligence.

Acknowledgement: Not Applicable.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present

study.

References

- [1] M. Rana, S. Srinivas, L.K. Jamili, I.A. Jaiswal, S. Nakka and S. Kasetti, "Real-Time Monitoring and Prediction of Blood Sugar Levels in Diabetic Patients with Functional Models," *In 2025 International Conference on Engineering, Technology & Management (ICETM)*, pp. 1-6, 2025.
- [2] I.A. Jaiswal, "Intelligent Cybersecurity Framework for Large-Scale RESTful Service Architectures," *International Journal of Research Radicals in Multidisciplinary Fields*, vol.2, no.1, pp.178–184, 2023.
- [3] I.A. Jaiswal, "Multilingual and culturally adaptive AI models for global education platforms," *International Journal for Research in Education (IJRE)*, vol.12, no.9, pp.17–27, 2023.
- [4] M. S. Prashanth, R. Aluvalu and M. V. V. Kantipudi, "Enhancing Health Product Traceability on the Blockchain: A Novel Approach for Supply Chain Management Inspection to AI," *EAI Endorsed Transactions on Pervasive Health & Technology*, vol. 10, no. 1, 2024.
- [5] H. K. Jani, M. V. V. Prasad Kantipudi, G. Nagababu, D. Prajapati and S. S. Kachhwaha, "Simultaneity of Wind and Solar Energy: A Spatio-Temporal Analysis to Delineate the Plausible Regions to Harness," *Sustainable Energy Technologies and Assessments*, vol. 53, 2022.
- [6] I. A. Jaiswal, "Natural language processing for security policy and log analysis," *International Journal of Research in All Subjects in Multi Languages (IJRSMML)*, vol.10, no.4, pp.57–67, 2022.
- [7] S. Khan, "Sales force security compliance: An in-depth study of GDPR, HIPAA, and PCI-DSS enforcement in cloud-based CRM systems and their implications for global enterprises," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol.8, no.9, pp.2211–2219, 2019.
- [8] S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in *Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)*, pp. 1–6, 2024.
- [9] S. Rani, D. Ghai and S. Kumar, "Reconstruction of wire frame model of complex images using syntactic pattern recognition," in *IET Conference Proceedings CP791*, vol. 2021, no. 11, pp. 8–13, 2021.
- [10] Swathi, S. Kumar, S. Rani, A. Jain, and R. K. M. V. N. M., "Emotion classification using feature extraction of facial expression," in *Proceedings of the 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, pp. 283–288, 2022.
- [11] S. Choudhary, S. Kumar, S. Gowroju, M. Gulhane, and R. Sri Lakshmi, Eds., "Genomics at the Nexus of AI, Computer Vision, and Machine Learning," *Hoboken, NJ, USA: John Wiley & Sons*, 2024.
- [12] S. Kumar, S. Choudhary, S. Gowroju, and A. Bhola, "Convolutional neural network approach for multimodal biometric recognition system for banking sector on fusion of face and finger," in *Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision*, pp. 251–267, 2023.
- [13] S. Choudhary, S. Kumar, M. Gulhane and M. Kumar, "Secured automated certificate creation based on multimodal biometric verification," in *Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision*, pp. 269–281, 2023.
- [14] I. A. Jaiswal, "AI-powered observability and incident prediction in distributed enterprise platforms," *Scientific Journal of Artificial Intelligence and Blockchain Technologies*, vol.1, no.1, pp.1–14, 2024.
- [15] S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in *Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)*, pp. 1–6, 2024.

-
- [16] I. A. Jaiswal, "High-Performance AI-Augmented Content Management Systems for Distributed Clouds," *International Journal of Multidisciplinary Innovation and Research Methodology*, vol.2, no.2, pp.90–97, 2023.
 - [17] S. Khan, "The role of zero-trust models in database security: Eliminating implicit trust and enforcing continuous verification in enterprise data access systems," *International Journal of Research in Electronics and Computer Engineering*, vol.6, no.1, pp.1622–1628, 2018.
 - [18] I. A. Jaiswal, "Self-healing REST services using artificial intelligence in multi-cloud environments," *Scientific Journal of Artificial Intelligence and Blockchain Technologies*, vol.1, no.3, pp.1–7, 2024.
 - [19] V. Saini, A. Jain, Anurag and M. V. V. Prasad Kantipudi, "An Efficient Approach for Improving the Performance of Autonomous Vehicle Using Advanced Computer Vision," in *International Conference on Soft Computing and Pattern Recognition*, pp. 164–174, 2023.
 - [20] S. Khan, "The role of privacy-preserving techniques in database security: Secure multi-party computation, differential privacy, and homomorphic encryption," *The Research Journal*, vol.3, no.4, pp.29–35, 2017.
 - [21] Segun KehindeSegun Kehinde, "AI in Everything, and Everything in AI: A Review of the Ubiquitous Role of Artificial Intelligence in Shaping the Next Technological Epoch," *JCAI*, vol. 3, no. 5, pp. 17–53, 2025.
 - [22] K. Cherladine, B.R. Rallabandi, A. Goel, K. Kaushik, A. Dhillon and M. Soni, "Generative adversarial defense models for simulated cyberattack scenarios in virtual networks," In *2025 International Conference on Digital Innovations for Sustainable Solutions (ICDISS)*, pp. 1–6, 2025.
 - [23] B.R. Rallabandi, "MEC-native 5G systems: Orchestration algorithms for ultra-low latency cloud-edge integration," *International Journal of Intelligent Systems and Applications in Engineering*, vol.8, no.4, pp.398–408, 2020.
 - [24] B. R. Rallabandi, "Joint deployment and operational energy optimization in heterogeneous cellular networks under traffic variability," *International Journal of Communication Networks and Information Security*, vol.10, no.3, pp.638–647, 2018.
 - [25] C. H. Neetha and M. P. Kantipudi, "Adaptive Edge-Based Bilinear Interpolation for Smart Healthcare," in *IET Conference Proceedings CP824*, vol. 2022, no. 26, 2022.
 - [26] R. Aluvalu, R. Venkatachalam, V. Uma Maheswari, R. J. Anandhi, M. V. V. Kantipudi, and S. Prashanth Mallellu, "IWHO-Based Cluster Head Selection for Vehicle-to-Vehicle Communication in Intelligent Transport System," *International Journal of Transport Development and Integration*, vol. 8, no. 2, pp. 154–166, 2024
 - [27] S. Velamuri, S. K. Sudabattula, M. V. V. Prasad Kantipudi, and N. Prabakaran, "Q-Learning Based Commercial Electric Vehicles Scheduling in a Renewable Energy Dominant Distribution Systems," *Electric Power Components and Systems*, pp. 1–14, 2023
 - [28] Swara Snehit Patil, "Artificial Intelligence: A Way to Promote Innovation," *JSIHS*, vol. 2, no. 1, pp. 1–5, 2024.
 - [29] A.B. Desai, B. Rallabandi, K.K. Gattupalli and M.S. Medarametla, "Agentic AI for rural connectivity and spectrum-aware networks to power smart agriculture ecosystems," In *2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT)*, pp. 1–6, 2025.
 - [30] S. Goyal, B. Rallabandi, K.K. Gattupalli and M.S. Medarametla, "Digital twin-enabled context-aware networks: Enhancing smart grid resilience through predictive intelligence," In *2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT)*, pp. 1–6, 2025.
 - [31] N. Tripathi, K.K. Gattupalli, B. Rallabandi and M.S. Medarametla, "AI-native Cloud-RAN orchestration for enterprise private 5G using digital twin models," In *2025 1st IEEE Uttar Pradesh Section Women in Engineering International Conference on Electrical Electronics and Computer Engineering (UPWIECON)*, pp. 851–857, 2025.
 - [32] S. Goyal, K.K. Gattupalli, B. Rallabandi and M.S. Medarametla, "Integrating terrestrial and non-terrestrial networks for reliable rural coverage in agriculture and smart grids," In *2025 1st IEEE*
-

Uttar Pradesh Section Women in Engineering International Conference on Electrical Electronics and Computer Engineering (UPWIECON), pp. 846–850, 2025.

- [33] N. Pachauri, V. Suresh, M. V. V. Prasad Kantipudi, R. Alkanhel, and H. A. Abdallah, “Multi-Drug Scheduling for Chemotherapy Using Fractional Order Internal Model Controller,” *Mathematics*, vol. 11, no. 8, 2023.
- [34] M. V. V. Prasad Kantipudi, S. Vemuri, N. S. Pradeep Kumar, S. Sreenath Kashyap, and S. Eslamian, “Proposing Model for Water Quality Analysis Based on Hyperspectral Remote Sensor Data,” in *Handbook of Hydroinformatics*, Elsevier, pp. 317–324, 2023.