ISSN: **3139-0854**

AI and Epistemic Injustice: Algorithmic Bias and the Marginalization of Knowledge in the Digital Age

Jagmanjot Singh[#1]
Phd Scholar, Centre for International Politics, Organisation and Disarmament, Jawaharlal Nehru University, New Delhi.
1jagmansingh18@gmail.com

Abstract
The rapid integration of Artificial Intelligence (AI) into academic research and knowledge production has fundamentally reshaped how knowledge is generated, accessed, and validated. While AI-driven tools promise efficiency, scalability, and democratization of knowledge, they also raise critical epistemological concerns regarding bias, exclusion, and power asymmetries. This paper examines the phenomenon of epistemic injustice in AI-mediated knowledge systems. It draws on the conceptual framework of Miranda Fricker and situating it within broader debates on power and knowledge articulated by Michel Foucault. The study argues that AI systems, particularly large language models and algorithmic search engines, are not neutral intermediaries but are active participants in the knowledge production. Trained on historically skewed and predominantly Global North-centric datasets, these systems risk perpetuating testimonial injustice by privileging certain voices while marginalizing others, especially from the Global South and non-English linguistic communities. Also, hermeneutical injustice emerges when marginalized knowledge systems, such as indigenous or regional epistemologies, are inadequately represented or entirely absent from digital corpora. This limits their intelligibility within AI-generated outputs. This helps to identify patterns of omission, bias, and epistemic hierarchy. The paper concludes that while AI has the potential to democratize access to knowledge, its current deployment risks reinforcing existing epistemic inequalities unless accompanied by deliberate policy interventions, inclusive data practices, and ethical governance frameworks. It calls for a rethinking of AI as not merely a technological tool but as a site of epistemic power that must be critically interrogated to ensure more just and inclusive knowledge production systems.

Keywords—Epistemic Inequality, Knowledge Hierarchies, Academic Integrity, Research Ethics, AI Regulation.

I. INTRODUCTION

For much of the past two decades, digital technologies were framed primarily as instruments of access i.e. tools that delivered existing knowledge more widely, quickly, and cheaply. The rise of large language models and AI-driven search and retrieval systems marks a decisive departure from this framing. AI is no longer merely a conduit through which knowledge travels. AI has become a gatekeeper that determines which knowledge is retrieved, how it is phrased, and which sources are treated as authoritative in the first place. Scholars increasingly describe this shift as a move from computational tools toward genuine epistemic actors, systems that co-produce meaning rather than simply transmit it (Palese, 2026). This repositioning of AI at the threshold of knowledge production, rather than at its periphery, is the starting point of this paper.

This shift brings into sharp notice a tension between the promises made on behalf of AI and the epistemological risks it actually carries. AI developers and enthusiasts have long promoted these systems as engines of efficiency and scale which are capable of democratizing access to information for populations historically excluded from formal knowledge institutions. Yet a growing body of scholarship complicates this optimistic narrative. Language models are demonstrably better at representing the small number of languages, and by extension the cultural and social worlds that dominate their training data. This leaves the majority of the world's linguistic communities thinly and inconsistently represented (Helm et al., 2023). Organizations working directly on digital inclusion have echoed this concern by noting that AI risks compounding an existing digital divide into a distinct and more consequential AI-divide along lines of region, language, and income (Gonzales, 2024). Far from resolving inequality, AI's efficiency gains may simply operate faster and at greater scale atop an already uneven epistemic terrain (Okolo, 2023).

Central to this paper is the claim that AI systems are not neutral intermediaries standing outside the knowledge they process, but active participants that shape, filter, and validate knowledge according to the patterns embedded in their training data and design choices. This is not a peripheral technical flaw correctable through minor adjustment. It reflects how power operates through systems that appear objective while quietly encoding particular ways of knowing as default and universal (Muzanenhamo & Power, 2024). Read through Miranda Fricker's (2007) account of epistemic injustice and Michel Foucault's (1980) analysis of the entanglement of power and knowledge, AI systems can be understood as sites where testimonial credibility is unevenly distributed and where entire interpretive frameworks may be rendered unintelligible. These effects are especially consequential for Global South and non-English-speaking knowledge communities.

The paper proceeds as follows. Section two develops the theoretical framework. Section three reviews existing literature on algorithmic bias and epistemic exclusion. Section four outlines the qualitative methodology used to examine AI-generated academic responses. Section five presents findings on patterns of omission and hierarchy. The paper concludes in section six. List of references is provided at the end of the paper.

II. THEORETICAL FRAMEWORK

This paper draws on two bodies of thought that, on the surface, occupy different disciplinary terrains but converge

ISSN: 3139-0854

productively once brought into conversation. Miranda Fricker's account of epistemic injustice supplies the analytical vocabulary needed to name specific harms that occur when someone is wronged in their capacity as a knower. Michel Foucault's writing on power and knowledge supplies the structural lens needed to explain why such harms are not incidental but built into the very systems that now mediate knowledge at scale. Taken together, these two frameworks allow the paper to move from description to diagnosis. Fricker tells us what is wrong and Foucault tells us why it keeps happening.

Fricker (2007) distinguishes two forms of epistemic injustice, and the distinction matters because each names a different mechanism of harm. Testimonial injustice occurs when a speaker's word is given less credibility than it warrants because of a prejudice the listener holds about the speaker's social identity. The wrong here is not a simple misjudgement of fact, it is a deflation of credibility that tracks group membership rather than the actual reliability of the testimony offered. A woman dismissed in a professional meeting, or an African economist's analysis discounted relative to an equivalent claim made by a European counterpart, are canonical examples of this pattern. Hermeneutical injustice operates at a different level. It occurs when a gap in the collective interpretive resources available to a society makes it difficult for someone to make sense of, or communicate, an experience that is nonetheless real to them. Fricker's own example, sexual harassment before the term existed to describe it, illustrates the point. The injustice lies not in a single hearer's prejudice but in a structural absence, a hole in the shared conceptual toolkit that a marginalized group had no equal part in building. Both forms of injustice, Fricker argues, are ultimately traceable to unequal participation in the practices that generate and circulate knowledge. This is the reason what makes her framework transferable to the study of AI systems that now perform exactly these functions of circulation and validation.

Recent scholarship on generative AI has already begun applying these categories directly. Muzanenhamo and Power (2024) show how large language models, by stripping away the provenance of the material they draw on, quietly reproduce a credibility hierarchy in which African intellectual contributions to fields such as accounting are absorbed without acknowledgment and returned to users in a form that reads as generically authoritative rather than situated and attributable, a mechanism that functions as testimonial injustice at industrial scale. Helm et al. (2023) demonstrate a parallel hermeneutical dynamic in the domain of language itself. Because AI language technologies are trained overwhelmingly on a small cluster of dominant languages, the concepts and social meanings embedded in underrepresented languages have no adequate route into the model's outputs. This leaves speakers of those languages structurally unable to have their experiences rendered intelligible through the tool millions of others now rely on for exactly that purpose.

Where Fricker offers categories of harm, Foucault offers an account of why those harms are systemic rather than accidental. His central claim, developed across several of his major works, is that power and knowledge are not separable phenomena that occasionally interact but are mutually constitutive. Every exercise of power carries with it a production of knowledge, and every regime of knowledge sustains a corresponding form of power (Foucault, 1980). Societies, on this view, operate according to a regime of truth, a set of institutional procedures that determine which statements count as true, who is authorized to make truth claims, and by what techniques those claims are verified and circulated. Crucially, Foucault insists this is not a matter of some statements being distorted by an otherwise neutral system. The system itself, including its scientific institutions, its media, and its technical apparatuses, is how truth gets produced in the first place. Applied to the present context, this reframes the question the paper is asking. An AI model is not a distorting mirror held up to some pre-existing body of neutral knowledge. It is itself a node within a contemporary regime of truth, one that decides, through training data curation, ranking, and output generation, which claims read as credible and which do not. Recent work extending Foucault's power/knowledge thesis directly to generative AI reaches a similar conclusion, arguing that platforms which mediate search, newsfeeds, and generated content now perform a gatekeeping function structurally comparable to earlier institutions that once held near-total authority over legitimate knowledge (Torkestani & Mansouri, 2025; Palese, 2026). Even organizations working on AI governance from a policy rather than philosophical standpoint have converged on this point, noting that the opacity of algorithmic systems around training data and ranking criteria makes independent verification of their truth-claims nearly impossible for the public they serve (Open Government Partnership, 2023).

The synthesis this paper relies on, works as follows. Fricker gives the diagnostic vocabulary. When an AI system's output disproportionately discredits certain sources or renders certain experiences unintelligible, we can name that harm precisely as testimonial or hermeneutical injustice rather than gesturing vaguely at bias. Foucault gives the structural explanation for why this happens with such consistency. AI systems function as regimes of truth in their own right, and the credibility hierarchies Fricker describes are not glitches within those systems but expressions of how the systems are constituted to produce authoritative knowledge in the first place. Reading Fricker through Foucault, and Foucault through Fricker, therefore does more than combine two theorists. It produces an analytical apparatus capable of identifying specific epistemic harms in AI outputs while simultaneously explaining their structural origin in how power operates through contemporary knowledge infrastructures.

III. LITERATURE REVIEW

Scholarship on bias in AI systems has, over the past several years, moved from broad observations about unfairness toward increasingly granular accounts of how training data itself encodes geographic and cultural asymmetries. Manvi et al. (2024) provide direct empirical confirmation of this pattern by showing that several widely deployed language models perform reliably on factual and reasoning tasks tied to wealthier, data-rich regions while degrading noticeably on equivalent tasks concerning the Global South. The authors trace back this disparity to the uneven geographic distribution of the web text these models are trained on. This is not an isolated finding. Related work on cultural and geopolitical bias has documented that model outputs tend to default to value systems associated with the countries where the leading AI firms are headquartered, embedding a particular cultural

ISSN: 3139-0854

standpoint as though it were a neutral baseline rather than one perspective among many. What this literature converges on, in short, is that dataset skew is neither marginal nor easily patched, it is a structural property of how these systems are built, since the corpora available at internet scale already reflect decades of unequal digitization, publishing, and archival investment favouring the Global North.

A second and more recent strand of literature engages directly with the question of epistemic injustice in algorithmic systems, extending Fricker's original framework, which was developed with human interlocutors in mind, into the domain of automated decision-making and text generation. Symons and Alvarado (2022) were among the first to argue systematically that algorithms themselves can be sources of epistemic injustice, given that they function as what Alvarado (2023) later terms epistemic technologies. These systems do not merely transmit information but actively curate, weight, and produce it. Their account of algorithmic opacity is particularly relevant here since it explains why testimonial and hermeneutical harms produced by AI are harder to contest than those produced by a biased individual. There is no single identifiable prejudice to appeal against, but only an inscrutable statistical process whose outputs nonetheless carry an unearned appearance of objectivity. Complementing this philosophical work, Mathiesen's (2015) earlier framework of informational justice, developed originally for library and information science, anticipated many of these concerns by insisting that access to information and the fairness of its organization are matters of justice in their own right, not merely technical or logistical questions. This claim has aged well as AI systems have taken over precisely the organizational functions Mathiesen had libraries in mind for.

A third body of work situates these epistemic concerns within the broader literature on the digital and linguistic divide. Helm et al. (2023) show that language modelling technologies remain concentrated around a small fraction of the world's languages. They produce what they term language modelling bias, a built-in representational advantage for politically and economically dominant languages. The World Bank's recent analysis of global AI exposure reaches a parallel conclusion from a development-economics angle. It notes that the usage and benefit of AI tools remain heavily skewed toward wealthier, more connected populations. This compounds rather than resolves existing digital inequality (World Bank, 2025). UNESCO-affiliated research has raised similar alarms by warning that an emerging AI divide risks entrenching disadvantage for already marginalized regions, languages, and communities well beyond the older problem of basic internet access (Gonzales, 2024).

Despite the density of this literature, a gap remains that this paper aims to address. Existing algorithmic epistemic injustice scholarship, exemplified by Symons and Alvarado (2022) and Alvarado (2023), tends to stay at the level of philosophical argument. It is rarely paired with structural accounts of power capable of explaining why these injustices recur so consistently across systems and firms. Conversely, the digital divide literature is empirically rich but rarely engages Fricker's distinctions with precision, and collapse testimonial and hermeneutical harms into a general notion of bias. Even recent comparative work such as Olaniyan et al. (2026), which interviews students across the Global North and South about their experience of generative AI, stops short of grounding its findings in an explicit power-theoretic account of why these disparities are structurally reproduced rather than incidental. This paper's contribution lies precisely in that combination: using Fricker's categories to diagnose specific harms in AI-generated academic content while using Foucault's account of power and knowledge to explain their structural persistence.

IV. METHODOLOGY

This study adopts a qualitative literature-based analytical design rather than primary data collection, since its central aim is theoretical synthesis rather than the empirical detection of bias. A rapidly growing body of scholarship already documents, through direct auditing and testing of AI systems, the patterns of omission, credibility hierarchy, and interpretive erasure that concern this paper (e.g., Muzanenhamo & Power, 2024; Helm et al., 2023; Kraft & Soulier, 2024; Mollema, 2025). Rather than replicating this empirical work, the paper's contribution lies in reading these existing findings systematically through Fricker's (2007) categories of testimonial and hermeneutical injustice and Foucault's (1980) account of power/knowledge.

Secondary sources were identified through targeted searches of academic databases (Scopus, Google Scholar) and relevant grey literature (UNESCO, World Bank, Brookings) using combinations of terms including epistemic injustice, algorithmic bias, AI and Global South, and language modelling bias. Sources were included where they provided empirical or documented evidence of AI-related knowledge exclusion, credibility distribution, or linguistic/cultural representation, with priority given to peer-reviewed and recent (2022–2026) publications. Each source was then read, following a directed approach to thematic analysis (Braun & Clarke, 2019), using Fricker's testimonial/hermeneutical distinction and Foucault's regime-of-truth concept as pre-established analytical categories rather than emergent codes.

The scope of this study is accordingly interpretive and diagnostic rather than exhaustive. It does not claim to survey all existing AI systems or languages, but to demonstrate, through representative secondary evidence, how a Fricker-Foucault framework illuminates the structural character of epistemic injustice in AI-mediated knowledge production.

V. ANALYSIS AND FINDINGS

The literature shows a recurring testimonial pattern. AI systems systematically privilege sources bearing the markers of Western academic authority such as journal indexing, institutional affiliation and English-language phrasing, while treating other forms of evidence as secondary. Muzanenhamo and Power (2024) illustrate this concretely in the accounting field by showing that ChatGPT absorbs African intellectual contributions into its outputs while stripping their provenance, so the resulting text reads as generically authoritative rather than attributable to a specific scholarly tradition. This is testimonial injustice in Fricker's precise sense. Credibility is not withheld outright but is quietly redistributed, detached from the identity of its original source and reassigned to the model's own authoritative voice. Kraft and Soulier (2024) extend this point and argue that knowledge-enhanced language models inherit an assumption that knowledge is a neutral category, which leads them to under-credit standpoint-based or situated claims relative to claims framed in the register of institutional

ISSN: **3139-0854**

science. De Proost and Pozzi (2023) add a further wrinkle. They argue that conversational AI is trained to sound fluent and confident regardless of source quality, so it can flatten the very credibility gradient Fricker's framework depends on. This makes the discounting of marginalized testimony harder for an ordinary user to detect than in human-authored text.

A second, arguably more corrosive pattern concerns hermeneutical injustice, the absence of interpretive resources needed to render an experience intelligible at all. Here the clearest example is conceptual flattening. Mollema (2025) documents cases where generative systems, faced with concepts native to non-Western epistemologies, do not simply omit them but manufacture plausible-sounding yet hollow substitutes, a phenomenon he terms hermeneutical erasure. A widely cited illustration of this dynamic is the Southern African concept of Ubuntu, whose communal, relational logic resists propositional translation but is nonetheless routinely rendered by AI systems as a loose synonym for collectivism, stripped of its ethical and cosmological specificity. The effect, as Mollema (2025) notes, is more insidious than a refusal to answer, since it disguises interpretive absence as competence.

These two patterns converge along a Global North/South axis. Language-gap research from Cohere (2024) finds that model performance and knowledge coverage degrade sharply outside a small cluster of high-resource languages, consistent with Joshi et al.'s (2020) earlier finding that a handful of languages account for the overwhelming share of NLP training data. Mim et al. (2024) document a parallel dynamic in generative image tools used in South Asian creative practice, where outputs default to Global North visual conventions unless users invest considerable extra effort steering the system otherwise. Khan (2025) frames this broader pattern as a form of AI colonialism, in which data and labour flow outward from the Global South while interpretive defaults remain concentrated in the Global North.

Read together through Fricker and Foucault, these findings suggest that testimonial and hermeneutical injustice in AI systems are not isolated technical glitches but two expressions of the same underlying asymmetry. A knowledge infrastructure built predominantly from Global North, English-language material will, automatically, credit that material more readily and render everything outside it less intelligible.

These patterns point toward concrete policy directions rather than purely diagnostic ones. Regulatory and institutional frameworks should mandate independent, third-party auditing of large language models' training data composition and output performance across languages and regions, following established auditing protocols (Landers & Behrend, 2023; Mökander et al., 2024), paired with transparency requirements comparable to those already proposed for algorithmic decision-making more broadly (Open Government Partnership, 2023). Equally important is direct investment in the data and interpretive infrastructure of underrepresented knowledge systems, not merely translating existing Global North corpora but resourcing the documentation of indigenous, regional, and non-English scholarship on its own terms, so that hermeneutical resources exist for AI systems to draw on rather than flatten. Academic institutions and publishers, meanwhile, should treat AI-assisted knowledge production as a site requiring the same epistemic scrutiny as citation practice or peer review, rather than folding it into generic responsible AI guidance (Higgins, 2024; Bhimani et al., 2026).

VI. CONCLUSION

This paper set out to examine whether AI systems function as neutral conduits of knowledge or as active participants in its production. The analysis points firmly toward the latter. Read through Fricker's (2007) account of testimonial and hermeneutical injustice and Foucault's (1980) understanding of power and knowledge as mutually constitutive, AI-mediated knowledge systems emerge not as background infrastructure but as a site of epistemic power in their own right, one that decides, often invisibly, whose testimony carries weight and whose interpretive frameworks are legible at all. The recurring privileging of Global North, English-language, institutionally credentialed sources, and the corresponding thinning or erasure of indigenous, regional, and non-English epistemologies, are not incidental defects awaiting a technical patch. They are the predictable output of a regime of truth built on an unrepresentative corpus and then deployed as though it were universal (Palese, 2026).

The central implication is that AI cannot be approached with the same unreflective trust often extended to a reference tool or a search index. Treating these systems as neutral, as much current academic and institutional practice still does, forecloses the very scrutiny needed to notice whose knowledge they are quietly displacing (Khan, 2025). What this paper calls for instead is sustained critical interrogation of AI as an epistemic actor, embedded within research methodology, publishing norms, and pedagogy rather than confined to a separate ethics conversation, so that convenience does not become a substitute for epistemic accountability. Recent scoping work on generative AI in scholarly publishing has already flagged decolonial analysis of algorithmic bias as an underdeveloped research agenda (Raitskaya & Tikhonova, 2025). This paper suggests that future research should extend that agenda specifically into disciplinary knowledge production in the Global South, tracking longitudinally whether the interventions proposed here actually shift what these systems come to know, and whose knowledge they continue to leave out.

REFERENCES

[1] Alvarado, R. (2023). AI as an epistemic technology. Science and Engineering Ethics, 29(5), 1–30. https://doi.org/10.1007/s11948-023-00451-3

[2] Bhimani, A., et al. (2026). Academic publishing in the age of generative AI. Business & Information Systems Engineering. https://doi.org/10.1007/s12599-026-01013-0

[3] Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[4] Cohere. (2024). The AI language gap. Cohere Research. https://cohere.com/research/papers/the-ai-language-gap.pdf

[5] De Costa, P. I. (2025). Learning to use GenAI for scholarly publication: Possibilities and pitfalls from an Indonesian perspective. Journal of English for Research Publication Purposes, 6(2), 290–315.

[6] De Proost, M., & Pozzi, G. (2023). Conversational artificial intelligence and the potential for epistemic injustice. The American Journal of Bioethics, 23(5), 51–53. https://doi.org/10.1080/15265161.2023.2191020

[7] Foucault, M. (1980). Power/knowledge: Selected interviews and other writings 1972–1977 (C. Gordon, Ed.). Pantheon Books.

ISSN: **3139-0854**

[8] Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.

[9] Gonzales, S. (2024, August 6). AI literacy and the new digital divide: A global call for action. Global AI Ethics and Governance Observatory, UNESCO. https://www.unesco.org/ethics-ai/en/articles/ai-literacy-and-new-digital-divide-global-call-action

[10] Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2023). Diversity and language technology: How language modeling bias causes epistemic injustice. Ethics and Information Technology, 25(3). https://doi.org/10.1007/s10676-023-09742-6

[11] Higgins, C. (2024). Toward un-WEIRDing academic publishing about language. International Journal of the Sociology of Language, 2024(289–290), 27–31.

[12] Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 6282–6293). Association for Computational Linguistics.

[13] Khan, R. (2025, July 14). From AI colonialism to co-creation: Bridging the global AI divide. Media@LSE Blog. https://blogs.lse.ac.uk/medialse/2025/07/14/from-ai-colonialism-to-co-creation-bridging-the-global-ai-divide/

[14] Kraft, A., & Soulier, E. (2024). Knowledge-enhanced language models are not bias-proof: Situated knowledge and epistemic injustice in AI. In Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3630106.3658981

[15] Landers, R. N., & Behrend, T. S. (2023). Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models. American Psychologist, 78(1), 36–49. https://doi.org/10.1037/amp0000972

[16] LSE Impact Blog. (2026, June 23). AI promised to democratise academic publishing – the evidence says otherwise. London School of Economics. https://blogs.lse.ac.uk/impactofsocialsciences/2026/06/23/ai-promised-to-democratise-academic-publishing-the-evidence-says-otherwise/

[17] Manvi, R., Khanna, S., Burke, M., & Ermon, S. (2024). Large language models are geographically biased. arXiv. https://doi.org/10.48550/arXiv.2402.02680

[18] Mathiesen, K. (2015). Informational justice: A conceptual framework for social justice in library and information services. Library Trends, 64 (2), 198–225.

[19] Mim, N. J., Nandi, D., Khan, S. S., Dey, A., & Ahmed, S. I. (2024). In-between visuals and visible: The impacts of text-to-image generative AI tools on digital image-making practices in the Global South. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Article 474). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642204

[20] Mökander, J., Schuett, J., Kirk, H. R., & Floridi, L. (2024). Auditing large language models: A three-layered approach. AI and Ethics, 4, 1085–1115. https://doi.org/10.1007/s43681-023-00289-2

[21] Mollema, W. (2025). A taxonomy of epistemic injustice in the context of AI and the case for generative hermeneutical erasure. AI and Ethics, 5(5), 5535–5555. https://doi.org/10.1007/s43681-025-00801-w

[22] Muzanenhamo, P., & Power, S. B. (2024). ChatGPT and accounting in African contexts: Amplifying epistemic injustice. Critical Perspectives on Accounting. https://doi.org/10.1016/j.cpa.2024.102735

[23] Okolo, C. T. (2023, June 1). AI in the Global South: Opportunities and challenges towards more inclusive governance. Brookings Institution. https://www.brookings.edu/articles/ai-in-the-global-south-opportunities-and-challenges-towards-more-inclusive-governance/

[24] Olaniyan, Y. D., Martins, M. O., & Al Maqrashi, R. H. (2026). Generative artificial intelligence and epistemic (in)justice: Perspectives from higher education students in the Global North and South. Frontiers in Human Dynamics. https://doi.org/10.3389/fhumd.2026.1790324

[25] Open Government Partnership. (2023). State of the evidence: Algorithmic transparency. https://www.opengovpartnership.org/wp-content/uploads/2023/05/State-of-the-Evidence-Algorithmic-Transparency.pdf

[26] Palese, R. (2026). Artificial truth: Algorithmic power, epistemic authority, and the crisis of democratic knowledge. Societies, 16(3). https://doi.org/10.3390/soc16030102

[27] Raitskaya, L., & Tikhonova, E. (2025). The 2025 landscape of generative AI in scholarly writing and publishing: A scoping review of uses and ethical approaches. Journal of Language and Education, 10(2).

[28] Symons, J., & Alvarado, R. (2022). Epistemic injustice and data science technologies. Synthese, 200(2), 1–26. https://doi.org/10.1007/s11229-022-03631-z

[29] Torkestani, M. S., & Mansouri, T. (2025). Will power return to the clouds? From divine authority to GenAI authority. arXiv. https://doi.org/10.48550/arXiv.2512.03076

[30] World Bank. (2025). Inequalities in use of and exposure to artificial intelligence. Atlas of Global Development. https://data360.worldbank.org/en/atlas/artificial-intelligence/