International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

Epistemic Authority And Legitimacy In The Age Of Generative Ai: A Normative Framework For Evaluating Ai As A Knowledge-Formatter In Scientific Inquiry

Jazimatul Husna[1*], Imilia Ibrahim[2]

[1*]Departement of Informations and Culture, Vocational School, Universitas Diponegoro, Semarang, Indonesia.
jazimatulhusna@lecturer.undiup.ac.id
[2]Faculty of Information Management, Universiti Teknologi MARA (UiTM), Sarawak, Malaysia.
Corresponding Author: Jazimatul Husna

Abstract– The growing integration of generative Artificial Intelligence (AI) into scientific inquiry raises foundational questions about epistemic authority and justified epistemic dependence. Classical accounts—ranging from Goldman's reliabilism to Zagzebski's preemption theory—treat epistemic authority as a normative relation between agents characterized by intentionality, reason-responsiveness, and accountability. Drawing on critiques by McMyler and Jäger and Shackel, this paper argues that such agential and relational conditions exclude AI systems from possessing epistemic authority in any classical sense. Yet AI increasingly shapes evidential practices, hypothesis formation, and interpretive processes within contemporary science. To evaluate this non-agential epistemic influence, the paper develops the Epistemic Authority Analysis Framework (EAAF), which distinguishes epistemic authority from epistemic legitimacy. The EAAF identifies three normative conditions—Epistemic Transparency, Delegative Trust, and Normative Reflexivity—under which reliance on AI can generate justified belief without conferring authority on the system itself. By framing AI as a "knowledge-formatter," the paper demonstrates how epistemic responsibility can be preserved while integrating AI into scientific inquiry. The framework clarifies the epistemic status of AI and offers a principled foundation for human–machine epistemic collaboration in increasingly algorithmic epistemic environments.
Keywords: Epistemic Authority; Epistemic Legitimacy; Generative AI; Knowledge-Formatter; Trust and Delegation.

I. INTRODUCTION

The rapid integration of generative Artificial Intelligence (AI) into scientific inquiry is transforming the structure of epistemic dependence. Systems capable of synthesizing, summarizing, and reformatting scientific information increasingly shape how researchers interpret evidence and form justified beliefs. This development raises a central philosophical question: can a non-agential system lacking intentionality, metacognition, and accountability serve as a legitimate epistemic contributor within scientific inquiry?

Classical epistemology provides a rigorous framework for addressing this question. Goldman's (1986) (1986) reliabilism grounds epistemic authority in truth-conducive cognitive processes; Fricker (2007) (2007) emphasizes the moral–epistemic norms governing credibility; and Zagzebski's (2012) (Zagzebski, 2012) preemption theory argues that rational agents may delegate epistemic authority to others whose cognitive performance they judge superior to their own. These accounts converge on a structural insight: epistemic authority is an agential, reason-giving, and normatively accountable relation. Authority presupposes reflective trust, intentionality, and the ability to assume epistemic responsibility.

Recent philosophical work challenges the relevance of these assumptions in AI-mediated epistemic environments. Hauswald (2025) (2025) argues that AI systems increasingly function as de facto epistemic authorities—not because they possess agential authority, but because they structure information flow within epistemic communities. Boyd (2022) (2022) shows how digital infrastructures reshape practices of trust by bypassing traditional markers of expertise. Wheeler (2020) (2020) further demonstrates that machine-generated outputs cannot constitute classical testimony, revealing a tension between mechanical reliability and normative epistemic standards. Spivack & Gillis Jonk (2025) (2025) highlight AI's structural cognitive limits—particularly the absence of self-knowledge and justificatory awareness—reinforcing that AI exerts epistemic influence without satisfying the classical requirements for authority. Despite this, existing AI-ethics literature often focuses on issues such as opacity, bias, and fairness (Bender et al., 2021; Burrell, 2016; Mittelstadt et al., 2016) rather than on the deeper normative problem of epistemic legitimacy: under what conditions, if any, is reliance on AI epistemically justified? Ferrario et al. (2024) warn against conflating AI's performance superiority with epistemic authority, while Constantin & Grundmann (2020) (2020) and Keren (2014) (2014) show that preemption depends on normative conditions AI cannot satisfy.

At the same time, Seger & Hall (2022) argue that AI can nevertheless improve the justificatory position of users by reshaping the epistemic environment through aggregation and calibration. Beisbart & Räz (2022) Räz (2022) similarly emphasize that the epistemic evaluation of algorithmic systems requires conceptual tools distinct from those used for agential sources. These developments reveal a persistent conceptual gap: classical theories explain why AI cannot be an authority, but they do not explain when reliance on AI can still be epistemically legitimate.

To address this gap, this paper reconceptualizes generative AI as a knowledge-formatter—a system that structures, filters, and recombines information in ways that shape human reasoning without itself being a knowing subject. Building on classical epistemic authority theory while integrating contemporary AI epistemology, the paper proposes the Epistemic Authority Analysis Framework (EAAF). The EAAF distinguishes epistemic authority from epistemic legitimacy and identifies three normative conditions—Epistemic Transparency, Delegative Trust, and Normative Reflexivity—under which reliance on AI systems can generate justified belief without incorrectly ascribing authority to the system itself. The central thesis is thus twofold: (1) AI cannot possess epistemic authority in the classical sense, because authority requires normative agency; yet (2) reliance on AI can be epistemically legitimate when embedded within epistemic environments that satisfy specific normative constraints. By articulating these constraints, the EAAF provides a principled foundation for integrating AI into scientific inquiry while preserving the normative architecture of human epistemic responsibility.

II. CLASSICAL THEORIES OF EPISTEMIC AUTHORITY

Classical epistemology conceives epistemic authority as a normative relation in which one agent is entitled to guide another's belief formation. This relation is structured by three conceptual components that appear across the dominant accounts: the superior reliability of the authoritative agent, the socially mediated distribution of credibility, and the agential capacity to assume epistemic responsibility. Together, these components articulate the conditions under which a source may serve as a rationally preferred guide for an inquirer.

Reliabilist approaches, exemplified by Goldman (1986) (1986), ground epistemic authority in the comparative truth-conduciveness of belief-forming processes. A source functions as an authority when its cognitive methods are more reliable than those of the inquirer, thereby making deference a rational strategy for attaining accurate beliefs. While this model identifies reliability as a defining feature of epistemic authority, it treats authority primarily as a functional property of cognitive performance.

A complementary strand of analysis centers on credibility and the moral–epistemic norms that govern its distribution. Fricker (2007) (2007) characterizes epistemic authority in terms of socially situated credibility attributions, shaped by intellectual virtue and the ethical demands of testimonial justice. Within such practices, authority arises through the virtuous exercise of communicative competence and integrity, embedded in a community's normative expectations. Credibility, on this view, is not merely descriptive but also evaluative, reflecting the agent's position within a morally regulated epistemic community.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

The most systematic articulation of epistemic authority appears in Zagzebski's (2012) (2012) account of preemption. For Zagzebski, an agent may rationally defer to another when she judges—through reflective self-assessment—that the other's perspective is more likely to lead her to the truth than her own. Authority is, therefore, a normative power: the authoritative belief supplies a preemptive reason that replaces the agent's independent deliberation on the matter. This model presupposes an agent capable of intentionality, understanding, and reason-responsive self-governance, and it situates epistemic authority within relationships of reflective trust.

Subsequent work further clarifies these conceptual commitments. McMyler (2014) (2014) analyses epistemic authority as a distinctively interpersonal normative power to generate preemptive reasons for belief. He stresses that testimonial authority involves a transfer of epistemic responsibility, such that the speaker assumes responsibility for the conscientiousness of the hearer's belief. Jäger and Shackel (2025) (2025), in turn, critically examines preemptionist models and develops a guided inquiry conception according to which authorities support the inquirer's reasoning by helping her navigate the underlying structure of reasons. This account emphasizes the agential features of authority—possessing reasons, communicating them competently, and exercising epistemic virtues—which render authority both relational and normatively structured. The Bayesian analysis developed by Jäger and Shackel (2025) (2025) likewise presupposes a source whose credences can be incorporated into an agent's epistemic calculus, thereby reinforcing the notion that authority involves agents situated within normative relations of epistemic guidance.

The classical landscape is further refined by Lackey's (2008) (2008) hybrid theory of testimony, which distinguishes between the normative power characteristic of authority and the epistemic features of testimony that can transmit justification independently of the speaker's agential virtues. Lackey shows that testimony may justify belief even when the speaker lacks the qualities typically associated with authority, provided the testimonial environment is appropriately reliable. This distinction between agential authority and non-agential justificatory transmission introduces conceptual space for differentiating the normative status of an authoritative agent from the epistemic conditions under which belief can nevertheless be justified. Reliability, credibility, preemption, and guided inquiry articulate different dimensions of this status, while hybrid theories of testimony distinguish authority from more general epistemic dependence. This conceptual structure provides the theoretical foundation for subsequent analysis.

III. WHY AI CANNOT SATISFY CLASSICAL CONDITIONS OF EPISTEMIC AUTHORITY

Classical theories converge on the view that epistemic authority is an agential, normative, and relational status. Its defining feature is the power to generate preemptive reasons for belief—reasons that rationally replace or override the inquirer's independent deliberation. As Zagzebski (2012) (2012) argues, such preemption presupposes reflective trust: an agent must autonomously judge that another agent's epistemic perspective is superior and intentionally adopt that perspective as her own. This structure is intelligible only among beings capable of intentionality, conscientious reasoning, and normative answerability.

Recent refinements reinforce this agential conception. McMyler (2014) (2014) characterizes epistemic authority as a form of normative power rooted in interpersonal responsibility. Testimonial authority, he argues, involves a speaker who assumes responsibility for the conscientiousness of the hearer's belief—a structure that cannot be reduced to performance or reliability alone. Jäger and Shackel (2025) (2025) likewise demonstrates that authority relations require agents who possess reasons, communicate them intelligibly, and engage in reason-responsive guidance. His critique of preemptionism and development of the guided inquiry model further show that authority is a socially embedded practice involving mutual recognition and epistemic accountability.

These agential requirements are not optional. They are grounded in the broader tradition of epistemic responsibility articulated by Roberts & Woods (2007) (Roberts & Wood, 2007), Montmarquet (1992) (1992), and Sosa (2007) (2007), Within this tradition, epistemic responsibility is understood as conscientious regulation of one's cognitive life—deliberating for good reasons, responding appropriately to evidence, and being answerable for epistemic failures. Authority presupposes precisely these virtues: an authoritative source must act from epistemic competence, acknowledge its normative role, and be evaluable for the quality of its reasons. Artificial systems lack all such capacities, and therefore cannot satisfy the conditions required for genuine epistemic authority.

This stands in contrast to epistemic legitimacy, a broader category that concerns when reliance on a source results in justified belief. Lackey's (2008) (2008) hybrid theory demonstrates that testimony can transmit justification even when the speaker lacks epistemic

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

virtue, provided the testimonial environment is reliable. Legitimacy, unlike authority, does not require intentionality or responsibility; it requires that dependence be structured in ways conducive to justification. This distinction is crucial: authority concerns the normative status of the source; legitimacy concerns the epistemic status of the inquirer's reliance.

Artificial systems occupy this space of legitimacy rather than authority. As Hauswald (2025) (2025) argues, AI systems exert de facto epistemic influence by structuring the informational environment—ranking, filtering, and synthesizing data in ways that shape human inquiry. Yet these systems lack intentional states, commitments, or the capacity to assume responsibility. Wheeler (2020) (2020) demonstrates that machine outputs cannot constitute testimony, as testimony presupposes an agent who asserts with epistemic commitment. Similarly, Ferrario et al. (2024) (2024) show that even when AI systems exceed human performance, superiority cannot confer authority because authority requires epistemic virtues and normative responsiveness that machines cannot possess. Ross (2024) (2024) adds that opacity in AI systems further undermines the justificatory structure required for authority: without responsible epistemic agency, reliability alone is insufficient.

Seger & Hall ((2022) illustrate how reliance on AI may nonetheless improve epistemic justification when AI systems are embedded in reliable epistemic environments—through calibration, uncertainty modeling, and systematic filtering of evidence. This form of dependence mirrors Lackey's distinction between authority and legitimacy: a source may support justified belief without possessing the agential capacities that define authority.These analyses reveal a deep conceptual asymmetry. Classical theories show that epistemic authority requires intentionality, responsibility, and participation in interpersonal normative structures. AI systems lack all such features. To reconcile this tension, a normative framework must distinguish authority from legitimacy and articulate the conditions under which reliance on non-agential systems can be epistemically justified. This motivation underlies the Epistemic Authority Analysis Framework developed in the next section.

IV. THE EPISTEMIC AUTHORITY ANALYSIS FRAMEWORK (EAAF)

The preceding sections reveal a structural tension: classical theories of epistemic authority require agential, normative, and interpersonal properties that artificial systems lack, yet scientific inquiry increasingly depends on AI systems that shape the evidential and justificatory landscape in ways previously reserved for human authorities. To resolve this tension, the Epistemic Authority Analysis Framework (EAAF) provides a unified theoretical structure for assessing when reliance on a non-agential system—specifically generative AI functioning as a knowledge-formatter—can be epistemically legitimate even though the system cannot qualify as an epistemic authority in the classical sense.

At its core, the framework distinguishes authority from legitimacy: authority denotes a normative power to generate preemptive reasons for belief, whereas legitimacy concerns whether an agent's reliance on a source can yield justified belief under appropriate epistemic conditions. This distinction preserves the insights of classical theorists such as Zagzebski (2012) (2012), McMyler (2014) (2014), and Jäger (2025) (2025)--who show that authority requires intentionality, responsibility, and reason-giving—while accommodating the structural reliability emphasized in AI epistemology by Hauswald (2025) (2025), Seger & Hall (2022) (2022), Ferrario et al. (2024) (2024), Wheeler (2020) (2020), and Ross (2024) (2024). AI systems do not generate preemptive reasons, but they can still be integrated into epistemic practice in ways that support justified belief if specific normative conditions are met. The EAAF articulates these conditions through the concepts of epistemic transparency, delegative trust, and normative reflexivity.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

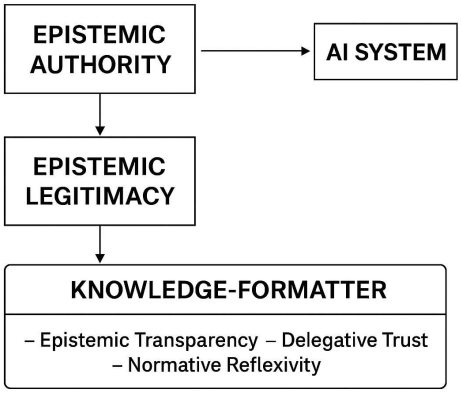

Figure 1. Conceptual Map of Authority, Legitimacy, and AI as a Knowledge-Formatter.

The Figure 1 visually positions generative AI outside the domain of classical epistemic authority while mapping three intersecting conditions—Epistemic Transparency, Delegative Trust, and Normative Reflexivity—that determine when reliance on AI is epistemically legitimate. Authority (left domain) is reserved for agential, reason-giving sources whose beliefs generate preemptive reasons; legitimacy (right domain) applies to systems lacking agency but capable of structurally enabling justified belief through reliable knowledge-formatting. The central triangular space indicates the normative zone where non-agential systems may be legitimately relied upon without being authorities.

The conceptual structure represented in Figure 1 clarifies the theoretical position of AI in epistemic practice. By locating AI outside the authority domain, the framework formally preserves the agential requirements defended in classical theories: preemption (Zagzebski), normative power (McMyler), guided inquiry and reason-responsiveness (Jäger and Shackel), and credibility-sensitive virtue (Fricker). At the same time, the diagram shows how AI may be incorporated into epistemic environments through structurally mediated pathways. Hauswald's analysis of AI as an "artificial epistemic authority" in a de facto sense is accommodated not by treating AI as an authority, but by situating AI within legitimacy conditions that regulate justified dependence. Likewise, Seger & Hall account of calibrated reliance, Wheeler's critique of machine testimony, and Ferrario's argument that AI cannot satisfy normative virtues all reinforce the boundaries indicated in the figure: AI systems lack authority but can still ground justified reliance when embedded within the appropriate epistemic infrastructure.

The first condition—Epistemic Transparency—requires that users understand not the internal mechanisms of AI systems, but the epistemic status, limitations, and operational constraints relevant to their outputs. This aligns with Ross's observation that opacity undermines justification unless users possess a clear understanding of an AI system's scope of competence. Transparency therefore concerns the intelligibility of epistemic risk, not technical explainability.

The second condition—Delegative Trust—requires users to form calibrated expectations about AI performance based on domain-specific reliability, verification practices, and the surrounding epistemic environment, echoing Seger & Hall argument that justified reliance depends on responsible calibration rather than blind trust.

The third condition—Normative Reflexivity—requires that scientists maintain active critical engagement with AI-produced representations: humans must remain epistemically answerable for the beliefs they form. In this respect, the framework extends Lackey's insight that justification can arise from environmentally reliable structures while preventing the erosion of agential responsibility that worries McMylera and Jäger & Shackel.

Through this synthesis, the EAAF resolves the conceptual tension between AI's non-agential nature and its increasing epistemic influence. It shows how AI can legitimately shape belief formation without ever becoming an epistemic authority. Authority remains the domain of normative agents; legitimacy is the domain of structured epistemic dependence. The framework thus provides a

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

principled method for evaluating AI's role in scientific inquiry: AI is not a knower but a knowledge-formatter, and its outputs can contribute to justified belief only when embedded within practices that satisfy the three conditions delineated above. By integrating classical epistemology with contemporary AI epistemics, the EAAF offers a rigorous foundation for assessing AI-mediated epistemic practices and clarifies the normative boundaries of responsible scientific reliance on non-agential systems.

V. APPLYING THE EAAF TO AI-MEDIATED SCIENTIFIC INQUIRY

The Epistemic Authority Analysis Framework (EAAF) offers a normative structure for evaluating when reliance on AI systems in scientific inquiry can be epistemically legitimate, even though such systems cannot meet the agential and interpersonal criteria for epistemic authority. While Section 2 demonstrated that AI cannot be an epistemic authority in the classical sense, contemporary research practices increasingly integrate AI as a structural contributor to epistemic processes. The question, therefore, is not whether AI can hold authority, but under what conditions its epistemic influence becomes legitimate.

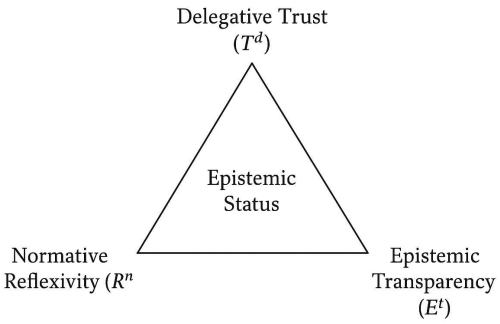

Figure 2. Applying Epistemic Authority Analysis Framework (EAAF) to AI-Mediated Scientific Inquiry.

Figure 2 conceptualizes legitimacy as emerging from the interaction of three conditions: Epistemic Transparency, Delegative Trust, and Normative Reflexivity. These conditions jointly determine the epistemic status of AI systems within inquiry. Each condition addresses a different dimension of the epistemic environment, and legitimacy is achieved only when all three operate in tandem.

Epistemic Transparency concerns the researcher's capacity to assess how AI outputs are generated, what evidential materials they rely on, and what epistemic limitations they entail. This requirement aligns with Ross's (2024) (2024) argument that opacity undermines justificatory structures and with Burrell's (2016) (2016) account of structural unintelligibility in machine learning systems. Transparency does not demand full interpretability but requires the availability of epistemic indicators—traceability, uncertainty estimates, error profiles—that enable researchers to evaluate AI outputs as evidential materials rather than opaque assertions. Delegative Trust involves the researcher's calibrated judgment that AI is the more reliable cognitive instrument for a specific domain or task. Seger & Hall (2022) (2022) emphasizes that justified reliance on AI requires calibration between user competence, system reliability, and task structure. This form of trust does not mimic authority or preemption; rather, it safeguards human epistemic autonomy by requiring agents to evaluate whether AI's contribution is appropriately aligned with their epistemic goals.

Delegative trust thus prevents the epistemic overreach identified by Ferrario et al. (2024) (2024), where AI is mistakenly treated as an authority solely based on performance advantages. Normative Reflexivity operates at the collective level. Scientific communities must institute norms that govern the use, oversight, and correction of AI systems. As Jäger and Shackel (2025) (2025) notes, epistemic authority among humans depends on socially embedded structures of responsibility and intellectual virtue; analogous structures are required to manage AI's epistemic role. Reflexive institutional practices—verification protocols, error audits, methodological guidelines—ensure that AI becomes a disciplined epistemic instrument embedded within the community's normative architecture.

International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119.

https://ijpsat.org/
Vol. 55 No. 1 January 2026, pp. 33-41

These conditions determine the epistemic status of AI systems. AI cannot be authoritative, because authority requires intentionality, responsibility, and reason-responsiveness. But AI can be epistemically legitimate when transparency enables evidential assessment, delegative trust calibrates reliance, and reflexive norms embed AI within communal epistemic practices. The EAAF thus provides a principled method for integrating AI into inquiry without collapsing the distinction between reliability, legitimacy, and authority. It resolves the conceptual tension identified in earlier sections by clarifying that while AI cannot possess authority, it can support justified belief formation under tightly constrained epistemic conditions.

VI. DISCUSSION

The analysis clarifies has advanced a fundamental reorientation: AI's epistemic role is not one of authority, but of legitimacy-dependent participation in scientific inquiry. This is not merely semantic distinction but a conceptual move with decisive theoretical consequences. McMyler (2014) (2014) and Jäger and Shackel (2025) (2025) show authority is necessarily relational, embedded in interpersonal normative structures AI cannot inhabit. The EAAF does not lament this; instead, it identifies the appropriate normative category: legitimacy. Following Lackey (2008), justified belief can arise from environmental reliability rather than agential power. This framework extends classical epistemology without abandoning its insights.

Yet why does EAAF succeed where alternative approaches falter? Three models attempted to handle AI's epistemic influence: Virtue-based approaches extend epistemic virtue epistemology directly to AI systems. However, virtues require intentional character development—commitment, habituation, evaluative stance toward truth-seeking—that AI execution fundamentally cannot instantiate. Virtue frameworks thus misconstrue what AI is. The EAAF avoids this by anchoring legitimacy in environmental structure, not agent virtue.

Competence-based models ground justification in domain-specific reliability. This captures something crucial: reliability matters, and domain-specificity is non-negotiable. Yet as Burrell (2016) and Ross (2024) demonstrate, reliability alone insufficient. High-competence systems pose epistemic dangers if opaque (deceiving users about actual confidence) or if users suffer deferential passivity. The EAAF incorporates competence via Delegative Trust, but adds Transparency (users understand limits) and Reflexivity (users maintain critical engagement). These additional conditions directly address opacity and passivity—dangers competence-based frameworks cannot recognize.

Reliabilist approaches treat AI as functionally equivalent to microscopes or statistical software. This appropriately denies AI agency. Yet it drastically undershoots AI's distinctive epistemic role. Microscopes are transparent media users operate; AI systems actively format evidence, curating interpretations and rendering certain conclusions plausible while marginalizing others. This recursive role over evidential landscapes itself requires normative accountability. The EAAF captures this distinctiveness through Normative Reflexivity: users must remain epistemically answerable for AI-shaped beliefs.

The EAAF's theoretical advantage lies in this systematic integration—incorporating reliability (Delegative Trust), competence-assessment (Transparency), and active user engagement (Reflexivity) while avoiding each alternative's categorical errors. It is superior not through arbitrary preference but through theoretical necessity.

This reorientation redefines scientific responsibility. Classically, responsibility locates in the authority—the person whose testimony one depends upon. The EAAF relocates responsibility to the inquirer—the scientist deciding when and how to rely on AI. This shift immediately reshapes scientific practice. Peer review must ask: Did the researcher satisfy legitimacy conditions? Was AI output transparent about limits? Did reliance calibrate to documented domain-specific performance? Did the researcher maintain critical engagement? Methodological guidelines must require explicit documentation of AI's epistemic boundaries, verification protocols, and researcher-performed critical assessment. Institutional design becomes epistemic design: scientific organizations must deliberately structure systems ensuring legitimacy conditions are met.

The three limitations warrant explicit acknowledgment. First, operationalization remains context-dependent: different scientific domains require different calibration standards. Second, the framework presupposes human reflexive capacity—an assumption failing in resource-constrained settings or among non-experts. Third, individual-reliance legitimacy differs from institutional-scale

AI integration; how EAAF scales remains unresolved. These are not failures but invitations for future philosophical and empirical work, establishing productive research horizons where this framework provides the necessary foundation.

VII. CONCLUSION

This paper has argued that generative AI challenges not the concept of epistemic authority itself but the epistemic environments in which authority and justification are negotiated. Classical theories—Zagzebski, Goldman, Fricker, McMyler, and Jäger—show that authority presupposes agential, reason-responsive, and normatively accountable capacities that AI systems lack. Consequently, AI cannot function as an epistemic authority in either the traditional or contemporary philosophical sense.

Yet AI's expanding role in scientific inquiry requires a normative account of when reliance on such systems remains epistemically justified. The Epistemic Authority Analysis Framework (EAAF) developed here provides this account by distinguishing epistemic authority from epistemic legitimacy. Legitimacy arises not from the properties of AI but from the conditions under which human agents and institutions structure their epistemic dependence on AI. Epistemic Transparency ensures that AI outputs remain evaluable; Delegative Trust preserves the agent's autonomy in calibrating reliance; and Normative Reflexivity embeds AI use within communal standards of oversight and accountability.

Under these conditions, AI can serve as an epistemically legitimate contributor to scientific inquiry without being misconstrued as an authority. The framework thus enables a principled integration of AI into contemporary epistemic practices while safeguarding the normative architecture that underwrites justified belief. More broadly, it reveals a shifting locus of epistemic responsibility—from agential authority to the design and governance of epistemic systems—offering a foundation for future philosophical work on human–machine epistemic relations.

REFERENCES

[1]. Beisbart, C., & Räz, T. (2022). Philosophy of science at sea: Clarifying the interpretability of machine learning. *Philosophy Compass*, *17*(6), e12830. https://doi.org/https://doi.org/10.1111/phc3.12830

[2]. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

[3]. Boyd, K. (2022). Trusting scientific experts in an online world. *Synthese*, *200*(1), 14. https://doi.org/10.1007/s11229-022-03592-3

[4]. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society*, *3*(1), 1–12. https://doi.org/10.1177/2053951715622512

[5]. Constantin, J., & Grundmann, T. (2020). Epistemic authority: preemption through source sensitive defeat. *Synthese*, *197*(9), 4109–4130. https://doi.org/10.1007/s11229-018-01923-x

[6]. Ferrario, A., Facchini, A., & Termine, A. (2024). Experts or Authorities? The Strange Case of the Presumed Epistemic Superiority of Artificial Intelligence Systems. *Minds and Machines*, *34*(3), 1–27. https://doi.org/10.1007/s11023-024-09681-1

[7]. Fricker, M. (2007). *Epistemic Injustice Power and the Ethics of Knowing*.

[8]. Goldman, A. I. (1986). Epistemology and cognition. In *Epistemology and cognition.* Harvard University Press.

[9]. Hauswald, R. (2025). Artificial Epistemic Authorities. *Social Epistemology*, *39*(6), 716–725. https://doi.org/10.1080/02691728.2025.2449602

[10]. Jäger, C., & Shackel, N. (2025). Testimonial Authority and Knowledge Transmission. *Social Epistemology*, *00*(00), 1–16. https://doi.org/10.1080/02691728.2025.2449608

[11]. Keren, A. (2014). Zagzebski on authority and pre-emption in the domain of belief. *European Journal for Philosophy of*

Religion, *6*(4), 61–76. https://doi.org/10.24204/ejpr.v6i4.145

[12]. Lackey, J. (2008). *Learning from Words: Testimony as a Source of Knowledge*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199219162.001.0001

[13]. McMyler, B. (2014). Epistemic Authority, Preemption and Normative Power. *European Journal for Philosophy of Religion*, *6*(4), 101–119. https://doi.org/10.24204/ejpr.v6i4.148

[14]. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

[15]. Roberts, R. C., & Wood, W. J. (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199283675.001.0001

[16]. Seger, E. A., & Hall, T. (2022). *a Comparative Investigation Into the Formation of Epistemically Justified Belief in Expert Testimony and in the Outputs of Ai-Enabled Expert Systems*. *May*.

[17]. Spivack, N., & Gillis Jonk, K. (2025). Epistemology and Metacognition in Artificial Intelligence: Defining, Classifying, and Governing the Limits of AI Knowledge. *Https://Www.Novaspivack.Com*, 1–33.

[18]. Wheeler, B. (2020). Reliabilism and the testimony of robots. *Techne: Research in Philosophy and Technology*, *24*(3), 332–356. https://doi.org/10.5840/techne202049123

[19]. Zagzebski, L. T. (2012). *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (Vol. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199936472.001.0001