

EPISTEMIC CERTAINTY AND THE LIMITS OF INFORMATIONAL KNOWLEDGE: A PHILOSOPHICAL INQUIRY

Hadeel Saed^{1*}

^{1*}Department of English Language and Translation, Faculty of Arts and Humanities.
Applied Science Private University. Amman, Jordan.
ORCID: 0000-0001-6833-2165
Corresponding Author Email: H_alsaed@asu.edu.jo

Abstract

Epistemic certainty has long been regarded as a central objective of philosophical inquiry, yet the expansion of digital information and artificial intelligence has renewed questions about whether absolute certainty is attainable. This paper examines the relationship between epistemic certainty and informational knowledge by integrating classical epistemology with contemporary theories of information and artificial intelligence. The manuscript argues that no finite informational agent can attain absolute certainty due to finitude, perspectivity, social mediation, and epistemic recursion, while clarifying its novel contribution beyond existing theories of epistemic limitations. Drawing upon the traditions of foundationalism, scepticism, fallibilism, and the epistemology of information, the study argues that these limitations are not temporary obstacles but intrinsic characteristics of finite cognitive systems. Rather than treating uncertainty as a weakness, the paper proposes that certainty should be understood as a regulative ideal that guides rational inquiry, critical evaluation, and continuous revision of beliefs. The analysis further explores the implications of this framework for artificial epistemic agents, particularly large language models, highlighting the importance of transparency, accountability, and epistemic responsibility in AI design and deployment. The paper concludes that reliable knowledge is best achieved through fallibilist inquiry, where intellectual humility and ongoing critical assessment replace the unattainable pursuit of infallible certainty.

Keywords: Epistemic certainty; Informational knowledge; Fallibilism; Epistemology of information; Artificial intelligence; Knowledge theory; Epistemic responsibility.

1. INTRODUCTION

The contemporary epistemic environment is characterized not merely by an abundance of information but by the increasing difficulty of distinguishing reliable knowledge from misleading or fabricated content. Digital platforms facilitate the rapid dissemination of misinformation and disinformation, while algorithmic ranking systems prioritize engagement over epistemic quality, creating filter bubbles and reinforcing confirmation biases. (Rowland, 2011; Sunstein, 2018). The emergence of synthetic media, including deepfakes and AI-generated text and images, further blurs the boundary between authentic and fabricated evidence, reducing trust in digital information ecosystems. (Chesney & Citron, 2019). At the same time, large language models can generate fluent, persuasive, yet factually inaccurate content, introducing new epistemic risks associated with automation, hallucination, and unverifiable knowledge claims. (Bender et al., 2021). Collectively, these developments create an epistemic environment in which greater access to information does not necessarily produce greater knowledge or certainty, but instead amplifies uncertainty regarding the reliability, provenance, and justification of beliefs.

The speed at which information can pass on the web, and the sheer volume of it, has been described as an epistemic paradox because an increase in the quantity of information flow can degrade the epistemic conditions of the internet. (Naser, 2025).

Epistemic Certainty and the Limits of Informational Knowledge: A Philosophical Inquiry

Despite the extensive philosophical literature on epistemic certainty, important theoretical gaps remain in explaining how classical epistemological concepts should be understood within contemporary information environments. Classical epistemology has traditionally examined certainty through foundationalism, scepticism, fallibilism, and theories of justification. At the same time, more recent scholarship has increasingly focused on the epistemology of information, digital misinformation, and the epistemic implications of artificial intelligence. Although extensive research has examined the nature and limits of epistemic certainty, these contributions remain largely fragmented across distinct research traditions. First, classical epistemology explains certainty in terms of justification, skepticism, knowledge, and epistemic regress, examining whether absolute knowledge is philosophically attainable (Sosa, 2021). Second, information epistemology investigates how digital information environments, misinformation, algorithmic mediation, and information overload influence knowledge acquisition and epistemic reliability (Floridi, 2009); (Levy, 2021). Third, AI epistemology examines the epistemic capabilities and limitations of artificial intelligence, including explainability, hallucinations, uncertainty, and machine reasoning (Bender et al., 2021); (Ji et al., 2023). However, these traditions generally address their respective domains in isolation and do not provide a unified theoretical framework explaining how common structural constraints simultaneously limit epistemic certainty in both human and artificial informational agents. Consequently, the underlying conceptual problem of whether shared structural conditions systematically prevent the attainment of absolute certainty across diverse epistemic systems remains unresolved. This study addresses this gap by proposing an integrated framework based on four structural constraints finitude, perspectivity, social mediation, and epistemic recursion that jointly explain why absolute certainty is unattainable for any finite informational agent.

This fragmentation has become increasingly significant in digitally mediated societies, where epistemic challenges arise not only from the quantity of available information but also from the structural characteristics of informational environments themselves. Existing studies have examined bounded rationality, testimonial knowledge, algorithmic bias, hallucination in large language models, and epistemic injustice as separate phenomena. However, comparatively little attention has been devoted to explaining how these seemingly distinct problems originate from common structural constraints that govern all finite informational agents. Consequently, the relationship between classical theories of certainty and contemporary informational epistemology remains conceptually underdeveloped, limiting our ability to understand why absolute epistemic certainty continues to elude both biological and artificial knowers.

This study develops an integrated conceptual framework showing that four structural conditions—finitude, perspectivity, social mediation, and epistemic recursion jointly constrain epistemic certainty across human and artificial informational agents. By unifying classical epistemology, information philosophy, and AI epistemology, it provides a common theoretical foundation for understanding the structural limits of knowledge.

This study investigates why absolute epistemic certainty remains unattainable for finite informational agents by addressing a clearly defined theoretical problem through explicit research questions, propositions, and an integrated conceptual framework grounded in four structural epistemic constraints. The main idea is that there does not exist any informational system (biological or artificial) that can be certain, for to be certain is to believe that the circumstances are fulfilled under which a finite informational agent can't fulfill. This claim does not entail scepticism. Instead, it inspires a fallibilist epistemology, a way of thinking about knowledge where the idea of certainty is a regulative ideal, an asymptote to guide inquiry without being realized in practice. The remainder of the paper is organized as follows. The paper is organized as follows. Sections 2–4 review classical epistemology, information philosophy, and AI epistemology. Section 5 introduces the proposed four-constraint framework. Section 6 evaluates its implications for fallibilism and competing theories; Section 7 applies it to artificial epistemic agents and large language models; and Section 8 discusses theoretical, practical, and future research implications.

2. THE CLASSICAL ARCHITECTURE OF EPISTEMIC CERTAINTY

2.1 Cartesian Foundationalism

René Descartes' method of hyperbolic doubt marks a defining moment in the history of epistemic certainty. Rather than claiming that whatever is doubted must be true, Descartes argues that the very act of doubting necessarily presupposes the existence of a thinking subject. Thus, the cogito (cogito, ergo sum) establishes not the certainty of external reality but the indubitability of the thinker's existence while thinking. (Paul, 2024). From this starting point, Descartes sought to reconstruct knowledge upon indubitable foundations capable of resisting sceptical doubt. His distinction between psychological certainty (subjective conviction) and epistemic certainty (objective impossibility of error) established a lasting framework for subsequent epistemological debates. However, the Cartesian project also exposes an enduring philosophical difficulty: although the cogito secures immediate self-awareness, it does not eliminate broader epistemic challenges concerning the interpretation, validation, and justification of knowledge claims about the external world. These unresolved limitations anticipate the structural constraints developed later in this paper. The proposed framework extends beyond Cartesian foundationalism by arguing that finitude constrains cognitive access to information, perspectivity limits viewpoint-neutral knowledge, social mediation shapes the validation of knowledge through epistemic communities, and epistemic recursion prevents the ultimate self-justification of epistemic standards. Descartes therefore provides the philosophical starting point for the pursuit of certainty, while simultaneously illustrating why absolute certainty cannot be generalized across finite informational systems.

2.2 Humean Scepticism and the Limits of Induction

David Hume fundamentally reshaped epistemology by arguing that empirical knowledge cannot be justified through deductive certainty. His critique of induction demonstrates that no finite number of observed instances can logically establish the truth of a universal proposition; rather, expectations about unobserved cases arise from habit and custom rather than rational necessity. (Calvente, 2022). Consequently, knowledge of matters of fact always remains provisional and open to revision, revealing an inherent limitation of finite epistemic agents. Hume extended this insight to testimonial knowledge by arguing that accepting another person's testimony ultimately depends upon inductive judgments regarding the general reliability of human communication. In this reductionist account, testimony derives its epistemic authority from prior experience showing that testimony is usually reliable rather than constituting an independent source of knowledge. Contemporary social epistemology has challenged this position through anti-reductionist theories, which maintain that testimony functions as a basic epistemic source under appropriate conditions. Nevertheless, both perspectives acknowledge that testimonial knowledge depends upon the integrity of informational transmission. This issue has become increasingly significant in contemporary digital environments, where misinformation, disinformation, synthetic media, and algorithmically amplified content undermine the reliability of testimonial networks. Hume's analysis therefore establishes the conceptual link between the problem of induction and the epistemology of testimony: because finite agents cannot directly verify most knowledge claims, they must rely upon socially transmitted information whose credibility always remains open to epistemic evaluation. This insight directly anticipates the framework developed later in this paper, particularly the structural constraints of social mediation and finitude, which explain why dependence on indirect information prevents the attainment of absolute epistemic certainty.

2.3 Moorean Certainty and the Common-Sense Tradition

G. E. Moore's responded to scepticism by arguing that certain common-sense beliefs—for example, that we have hands or that the external world exists—are more certain than the philosophical arguments intended to undermine them (Carroll, 2025). Rather than proving these beliefs through deductive reasoning, Moore treated them as more epistemically secure than sceptical conclusions.

Ludwig Wittgenstein extended this insight through the concept of *hinge propositions*, arguing that some fundamental commitments function as the unquestioned background that makes inquiry and evidence evaluation possible. However, hinge propositions do not establish absolute epistemic certainty because they are not propositions verified by evidence but practical presuppositions underlying epistemic practices. Consequently, they support rather than undermine the central argument of this paper. The proposed four-limit framework interprets such commitments as provisional conditions of inquiry that remain constrained by finitude, perspectivity, social mediation, and epistemic recursion. Contemporary challenges such as deepfakes and synthetic media (Habgood-Coote, 2023) further demonstrate that even these background assumptions are vulnerable within digitally mediated informational environments, reinforcing that certainty functions as a regulative condition of inquiry rather than an attainable epistemic state.

3. JUSTIFIED TRUE BELIEF AND ITS DISCONTENTS

3.1 The Tripartite Analysis

The analysis is a form of Cartesian infallibilism that qualifies it by replacing certainty with justification as the epistemic criterion. The JTB analysis can be understood from the point of view of the epistemology of information as demanding informational adequacy; belief-formation processes of the subject must be reasonably sensitive to the relevant information in the world. (Cabrera, 2023) Added to this intuition is distinguishing beliefs and the facts to which they correspond and insisting that the facts in question must be in a suitable causal relation to the belief.

3.2 The Gettier Problem

Edmund Gettier demonstrated that justified true belief is insufficient for knowledge by presenting cases in which a belief is both true and justified, yet true only because of epistemic luck rather than an appropriate connection between justification and truth. (Feldman, 2024). The justification supports the belief, but it reaches the truth through a defective evidential pathway. Subsequent responses, including no-false-lemmas, defeasibility, safety, and sensitivity theories (Paulson, 2024) attempt to explain how knowledge requires a reliable connection between evidence and truth. This insight is increasingly relevant in digital environments, where misinformation, synthetic media, and AI-generated content can produce accidentally true beliefs supported by misleading or unreliable evidence.

3.3 Reliabilism and Process Epistemology

Alvin Goldman's process reliabilism explains how reliable cognitive processes can generate justified knowledge (Goldman, 2025). However, the assessment of reliability itself depends upon finite evidence, situated perspectives, socially mediated evaluation, and recursively justified standards. Consequently, reliabilism complements but cannot replace the proposed four-limit framework explaining structural constraints on epistemic certainty. Reliabilism is an explicitly information-theoretic framework: that is, it regards the cognitive system as an information-processing system, and judges its output based on whether or not it correlates with truth. The generality problem, which describes why it is hard to specify which type of process is relevant for the evaluation of a particular belief, is a deeper issue in the case of artificial epistemic systems: specifying which process to use when evaluating the outputs of large language models whose inner workings are opaque.

4. THE EPISTEMOLOGY OF INFORMATION

4.1 Information, Content, and Truth

According to Luciano Floridi, semantic information consists of well-formed, meaningful, and truthful data, implying that misinformation does not constitute information in the epistemically relevant sense but instead represents defective informational content. (Floridi, 2013) This conception links the quality of information directly to the quality of knowledge. James Owen Weatherall (Habgood-Coote, 2023)

introduces the notion of an epistemic apocalypse to describe a condition in which synthetic media, deepfakes, and large-scale digital deception systematically undermine the evidential reliability of photographs, testimony, and other traditional sources of knowledge, thereby eroding justified belief formation. This concern complements an information-theoretic account of knowledge, which holds that knowledge depends upon the reliable transmission of information through causally appropriate channels. When these channels are degraded by noise, manipulation, or algorithmic distortion, epistemic reliability declines. Within the present framework, such failures illustrate the structural limits of social mediation and finitude, showing that reliable knowledge depends not only on information availability but also on the integrity of the informational systems through which knowledge is transmitted.

4.2 The Epistemology of Testimony

Testimony is one of the primary sources of human knowledge because individuals acquire most of what they know about history, science, politics, and everyday events through the reports of others rather than direct observation. Since testimonial knowledge depends on accepting another person's claims, its epistemic value rests upon the credibility assigned to the speaker and the perceived reliability of the communicative context. Contemporary social epistemology has shown that credibility assessments are not always based solely on evidence. These problems become more severe in digital environments, where bots, trolls, coordinated inauthentic accounts, and synthetic identities undermine the interpersonal trust on which testimonial knowledge depends (Kermani). Online manipulation therefore corrupts not only individual testimony but also the informational networks through which knowledge is socially transmitted. This progression demonstrates why social mediation constitutes a structural limit within the proposed framework: because finite informational agents necessarily depend on socially mediated testimony, distortions in credibility, communication, and information transmission inevitably constrain the attainment of absolute epistemic certainty.

4.3 Formal Epistemology and Probabilistic Warrant

Bayesian epistemology models rational belief through credence functions that are updated according to Bayesian conditionalization when new evidence is acquired. Under the standard Bayesian framework, assigning a proposition a credence of 1 represents full certainty and, subject to the regularity assumption and standard conditionalization, such a credence is not revised by subsequent evidence because conditionalization preserves certainty (Titelbaum, 2022). This interpretation should be understood as a normative ideal rather than a descriptive account of finite cognition. Titelbaum (2022) further argues that Bayesian rationality presupposes idealized conditions, including logical consistency and, in many formulations, logical omniscience. As (Landes, 2022) observes, these assumptions create a tension between normative Bayesian models and the computational limitations of finite epistemic agents, whether human or artificial. Within the present framework, Bayesian certainty therefore illustrates an idealized epistemic limit rather than an attainable cognitive state. The impossibility of satisfying these ideal conditions in finite informational systems reinforces the structural constraints of finitude and epistemic recursion, supporting the paper's central claim that absolute epistemic certainty remains unattainable.

4.4 Method of Conceptual Analysis and Theoretical Integration

This study employs a method of conceptual analysis and theoretical integration to develop a unified philosophical framework explaining why absolute epistemic certainty remains unattainable for finite informational agents. As a conceptual inquiry, the study does not analyse empirical data; instead, it systematically evaluates philosophical arguments, compares competing epistemological theories, and synthesizes concepts into a coherent explanatory model. The objective is not to determine which epistemological tradition is correct, but to identify structural features that recur across diverse accounts of knowledge and justification.

The analysis was organised around five philosophical traditions selected because each addresses a

fundamental dimension of epistemic certainty. Classical epistemology (Descartes, Hume, Moore, Wittgenstein, and Gettier) was examined to establish the historical foundations of certainty, scepticism, induction, common-sense knowledge, and justified true belief. Fallibilism and reliabilism (Peirce, Popper, Goldman, and Williamson) were included because they explain how reliable knowledge remains possible despite the continuing possibility of error. Bayesian epistemology (Titelbaum; Landes) was selected to analyse certainty as idealized credence under formal models of rational belief revision. Philosophy of information (Floridi) was incorporated to examine how information quality, semantic content, and reliable information transmission constrain knowledge acquisition. Finally, social epistemology and AI epistemology were included to analyse testimony, epistemic injustice, misinformation, algorithmic mediation, synthetic media, and the epistemic limitations of large language models in contemporary digital environments.

Across these traditions, the analysis compared four central epistemological concepts: (1) the nature of epistemic certainty, (2) conditions for knowledge justification, (3) the role of information and evidence in belief formation, and (4) the structural conditions governing epistemic reliability. Rather than evaluating these concepts independently, the study examined how each tradition explains the relationship between finite cognition, informational constraints, and the possibility of eliminating epistemic error. Areas of theoretical convergence and disagreement were identified through systematic comparison of their underlying assumptions, explanatory strengths, and unresolved conceptual limitations.

The proposed framework emerged through an iterative process of theoretical integration. Concepts recurring across multiple traditions were grouped according to their explanatory role rather than their historical origin. This comparative synthesis identified four structural limitations that consistently appeared despite differences in philosophical orientation: finitude, reflecting limits on cognitive and informational resources; perspectivity, reflecting the situated character of knowledge; social mediation, reflecting the dependence of knowledge on testimonial and institutional processes; and epistemic recursion, reflecting the regress involved in justifying epistemic standards themselves. These dimensions were retained because they repeatedly explained why certainty could not be fully secured across both human and artificial epistemic systems.

The resulting framework was evaluated against competing epistemological theories, including foundationalism, coherentism, reliabilism, Bayesian epistemology, and knowledge-first epistemology. The evaluation did not seek to replace these approaches but assessed whether each adequately explained the common structural conditions limiting certainty across biological and artificial informational agents. The proposed framework was considered theoretically successful to the extent that it incorporated the explanatory insights of these traditions while accounting for epistemic constraints that none of them addresses comprehensively in isolation. Its explanatory adequacy was further assessed by examining its capacity to interpret contemporary challenges associated with misinformation, synthetic media, digital platforms, and large language models within a single integrated philosophical account.

By explicitly specifying the selection of philosophical traditions, the concepts compared, the procedure used to identify recurring structural limitations, and the criteria employed to evaluate competing theories, this methodological approach enhances the transparency, analytical rigor, and reproducibility of the conceptual argument. The framework therefore emerges from a systematic process of philosophical comparison and theoretical integration rather than from unsupported conceptual assertion.

The paper develops an original conceptual framework comprising four interrelated structural limits.

Table 1: Analytical Development of the Four Structural Limits from the Literature

Literature Tradition	Main Epistemic Problem Identified	Structural Limitation Derived	Relevance to Human Knowledge	Relevance to AI Knowledge
----------------------	-----------------------------------	-------------------------------	------------------------------	---------------------------

Classical Epistemology (Descartes, Hume, Moore, Wittgenstein, Gettier)	Whether absolute certainty is possible; skepticism, induction, justification, and epistemic luck	Finitude and Epistemic Recursion	Human cognition cannot eliminate all possible error or completely justify its own standards of knowledge.	AI systems similarly operate with bounded information and cannot achieve complete self-validation of outputs.
Fallibilism and Reliabilism (Peirce, Popper, Goldman, Williamson)	Knowledge remains possible despite the continuing possibility of error; reliability replaces certainty	Finitude and Epistemic Recursion	Reliable knowledge depends on revisable methods rather than infallible certainty.	Model reliability depends on evaluation procedures that themselves require justification and continual revision.
Bayesian Epistemology (Titelbaum; Landes)	Rational belief is represented by credence updated through conditionalization under ideal assumptions.	Finitude	Finite agents cannot satisfy ideal Bayesian assumptions such as logical omniscience and perfect evidence processing.	AI systems approximate probabilistic reasoning but remain computationally bounded and resource constrained.
Philosophy of Information (Floridi)	Knowledge depends on truthful semantic information and reliable information transmission.	Finitude and Social Mediation	Knowledge is constrained by information quality, availability, and communication channels.	AI performance depends on the integrity of training data, information quality, and reliable transmission of content.
Social Epistemology (Harding, Catala, Habgood-Coote)	Testimony, credibility assessment, epistemic injustice, misinformation, and social validation	Social Mediation	Knowledge depends upon socially mediated testimony that may be distorted by power, bias, or misinformation.	AI systems inherit biases from socially produced data and may amplify misinformation and unequal credibility.
AI Epistemology (Bender et al.; Ji et al.; Floridi; Hauswald)	Hallucinations, explainability, synthetic media, benchmark limitations, and uncertainty	Perspectivity, Social Mediation, and Epistemic Recursion	Human users must critically evaluate AI-generated knowledge rather than treat it as authoritative.	AI systems possess bounded training perspectives, depend on socially generated data, and cannot fully justify their

				own epistemic standards.
Integrated Conceptual Synthesis (This Study)	Existing traditions explain individual limitations but not their common structural basis across human and artificial epistemic systems.	Finitude, Perspectivity, Social Mediation, and Epistemic Recursion	Provides a unified explanation for why finite human agents cannot attain absolute epistemic certainty.	Extends the same structural explanation to artificial epistemic agents and large language models, demonstrating common epistemic constraints.

5. THE FOUR LIMITS OF INFORMATIONAL KNOWLEDGE

The preceding analysis demonstrates that no single epistemological tradition provides a complete explanation of why absolute certainty remains unattainable for finite informational agents. Foundationalism identifies the aspiration for indubitable knowledge but struggles to establish universally immune foundations. Scepticism exposes the fragility of justification yet offers limited guidance for productive inquiry. Fallibilism acknowledges the revisability of knowledge without fully explaining the structural origins of epistemic limitation, while contemporary theories of information and artificial intelligence primarily address particular informational challenges rather than developing a comprehensive philosophical account of certainty. These observations suggest the need for an integrative framework capable of unifying insights that have largely remained dispersed across different philosophical traditions.

The framework comprises four dimensions because each corresponds to a distinct and indispensable stage of epistemic inquiry: information acquisition (finitude), interpretation (perspectivity), validation (social mediation), and justification (epistemic recursion). No dimension is reducible to another, and together they provide a comprehensive explanation of why finite informational agents cannot attain absolute certainty.

The first dimension, finitude, concerns the cognitive and computational limitations that restrict every epistemic agent's ability to acquire and process information. The second dimension, perspectivity, recognises that all knowledge is generated from situated cognitive and social viewpoints, preventing access to an entirely perspective-independent understanding of reality. The third dimension, social mediation, acknowledges that knowledge is fundamentally produced, transmitted, evaluated, and corrected through social institutions, communicative practices, and informational networks rather than by isolated individuals. Finally, epistemic recursion captures the unavoidable regress involved in justifying the standards by which knowledge itself is evaluated, demonstrating that no epistemic criterion can establish its own authority without further justification.

These four dimensions should therefore be understood not as independent limitations but as mutually reinforcing structural properties of finite epistemic systems. Finitude constrains the information available to an agent; perspectivity shapes how that information is interpreted; social mediation governs how informational claims are validated within epistemic communities; and epistemic recursion ensures that every justificatory framework remains itself open to critical examination. Their interaction provides a unified philosophical explanation for why certainty cannot be achieved through the accumulation of additional information alone. Even unlimited computational power or vastly expanded datasets cannot eliminate these structural characteristics because they arise from the architecture of finite cognition itself rather than from temporary informational deficiencies.

Accordingly, the framework developed here constitutes the principal theoretical contribution of this study. Rather than viewing uncertainty as an accidental consequence of incomplete information, it

reconceptualises uncertainty as an inherent feature of finite informational systems. This shift moves the philosophical debate beyond isolated discussions of scepticism, justification, or artificial intelligence by providing an integrated explanatory model that applies equally to biological knowers and contemporary artificial epistemic agents.

5.1 Finitude

The first structural dimension of the proposed framework is finitude, referring to the intrinsic cognitive and computational limits that characterize every finite informational agent. Human knowers possess bounded memory, attention, reasoning capacity, and time, while artificial systems operate under constraints imposed by model architecture, training objectives, finite datasets, inference procedures, retrieval quality, context windows, tool access, and computational resources (Ji et al., 2023; Bender et al., 2021). Although large language models can generalize beyond individual training examples by learning statistical regularities and producing novel combinations, their outputs remain conditioned by these underlying constraints rather than reflecting unrestricted epistemic access. Building on Herbert A. Simon's theory of bounded rationality (Stein) the present framework argues that finite agents necessarily rely on heuristics and incomplete evidence rather than exhaustive evaluation of all relevant possibilities. This claim should not be interpreted as asserting that every philosophical conception of certainty requires complete defeater elimination. Cartesian self-knowledge, mathematical proof, and Wittgensteinian hinge propositions represent domain-specific exceptions. Rather, the argument concerns absolute empirical epistemic certainty, where infallibilist accounts require the exclusion of undefeated defeaters. Because finite informational agents cannot guarantee that every potentially relevant defeater has been identified or assessed, certainty remains structurally limited. Increasing computational power or information availability may improve epistemic reliability but cannot eliminate the constitutive constraints imposed by finite cognition, thereby making finitude a foundational structural limit for both human and artificial epistemic systems.

Table 2: Conceptual Distinction between the Four Structural Limits

Structural Limit	Definition	Primary Mechanism	Type of Epistemic Vulnerability	Difference from Other Limits	Human Example	AI Example
Finitude	The intrinsic limits on cognitive and computational resources available to any finite epistemic agent.	Bounded memory, processing capacity, time, available evidence, and computational resources.	Incomplete evidence and inability to evaluate every relevant possibility.	Concerns resource limitations, not interpretation, validation, or justification.	Scientists cannot examine all possible evidence before reaching conclusions.	LLMs are constrained by training data, model architecture, objective functions, inference procedures, retrieval quality, context windows, tool access, and computational resources.
Perspectivity	Knowledge is	Interpretive framing and	Partial interpretation	Concerns how	Two historians	AI reflects linguistic,

Epistemic Certainty and the Limits of Informational Knowledge: A Philosophical Inquiry

	necessarily acquired and interpreted from a situated cognitive, historical, cultural, or theoretical standpoint.	standpoint dependence.	n and viewpoint-dependent understanding.	evidence is interpreted, not whether evidence exists or is socially validated.	interpret identical archival evidence differently.	cultural, and representational biases embedded within training data and model design.
Social Mediation	Knowledge is generated, transmitted, validated, and revised through social and institutional processes.	Testimony, peer review, institutions, communication networks, and credibility allocation.	Distortion through misinformation, bias, unequal credibility, and institutional failure.	Concerns how knowledge circulates and is validated socially, rather than individual interpretation.	Public trust depends on expert testimony and scientific institutions.	AI outputs inherit biases from socially produced datasets and may amplify misinformation through digital platforms.
Epistemic Recursion	Every standard used to justify knowledge itself requires further justification.	Recursive evaluation of epistemic standards.	Infinite regress or provisional stopping points in justification.	Concerns justification of epistemic standards, not cognitive resources or interpretation.	Philosophers debate why one method of justification should itself be trusted.	AI benchmark evaluation depends on metrics whose validity also requires independent justification.

5.2 Perspectivity

The perspectivity constitutes the second structural limitation because all knowledge is acquired and interpreted from a particular cognitive, historical, cultural, or theoretical standpoint. This claim should not be understood as endorsing epistemic relativism or implying that all beliefs are equally justified. Rather, it recognises that objectivity is achieved through the critical comparison of multiple perspectives rather than through a view from nowhere. Sandra Harding's standpoint epistemology argues that different social locations disclose distinct epistemic resources, making diverse perspectives essential for identifying otherwise overlooked assumptions and biases. (Harding, 1986). Consequently, critical pluralism, intersubjective scrutiny, peer criticism, and methodological correction improve the reliability and objectivity of knowledge by exposing errors that remain invisible within any single perspective. However, these corrective practices reduce rather than eliminate epistemic uncertainty because no finite informational agent can occupy every relevant standpoint or anticipate every potential defeater. In artificial intelligence, perspectivity appears through biases embedded in training data, model architecture, and design choices, reflecting uneven linguistic, cultural, and geographical representations (Spivack, 2025). Perspectivity therefore constrains absolute certainty while remaining fully compatible with the pursuit of increasingly

objective and reliable knowledge.

5.3 Social Mediation

The third structural constraint is social mediation since the majority of human knowledge is acquired, transmitted, validated, and corrected by social and institutional processes, but not only by direct individual observation. This does not mean to say that knowledge is exclusively a social product or that the individual's cognitive process and the evidence itself are unimportant. Instead, knowledge is produced through the interplay of individual thinking, evidence, and collective epistemic practices. Science communities have been shown by (Kuhn, 1997) to be central to the construction and reconstruction of the standards for evaluating knowledge claims, and today, social epistemology focuses on the strength of objectivity that comes from peer criticism, intersubjective scrutiny, and institutional accountability rather than individual cognition. However, there is a chance of distortion in the transmission of information. What and who is believed and how knowledge flows is shaped by credibility judgments, power dynamics, misinformation, and epistemic injustice (Kaya & Benevento, 2022). In digital spaces, the vulnerabilities are exacerbated by the algorithms of social media that favour engagement over epistemic reliability (Kirmayer, 2024), and by the use of large language models that can exacerbate socially-produced biases while seemingly acting as neutral epistemic authorities (Hauswald, 2025). Therefore, the social mediation contributes to the epistemic reliability of the collective evaluation process, but also brings in structural vulnerability, which prevents certainty,

5.4 Epistemic Recursion

Epistemic recursion is the fourth structural constraint since each justification for a claim to knowledge must be based on standards whose validity must be justified. This is the classical problem which Agrippa formulated in the form of his trilemma, to which the three responses—foundationalism, coherentism, and contextualism are given without doing away with recursive justification. In the modern example, it is used in the evaluation of AI. Benchmarks are used to model the performance, but the benchmarks rely on a set of justifiable choices of datasets, human labelling procedures, scoring rules, evaluation measures, and assumptions of task success. Further epistemic justification is needed for each of the evaluative standards. Epistemic recursion is not the end of inquiry, but an illumination of the fact that any kind of epistemic system, whether human or artificial, is in need of provisional, revisable justificatory standards, and that rationality is structurally impossible without provisionality, but possible with it.

5.5 Integrating the Four Structural Limits

Although each of the four structural limits has been examined individually, their explanatory power lies primarily in their interaction rather than in their independent effects. The proposed framework should therefore be understood as an integrated model of finite epistemic systems in which each limitation reinforces the others throughout the process of knowledge acquisition, evaluation, and revision. None of the four dimensions operates in isolation, and attempts to overcome one limitation invariably reveal the continuing influence of the remaining constraints.

The epistemic process begins with finitude, which limits the quantity of information that any cognitive agent can acquire, process, or retain. Because finite informational agents cannot exhaust all relevant evidence or consider every possible alternative, their knowledge claims necessarily depend upon incomplete informational resources. This incompleteness, however, is not merely quantitative. The selection, interpretation, and organisation of available information are inevitably influenced by perspectivity, whereby every epistemic agent approaches reality from historically, socially, culturally, and cognitively situated standpoints. Consequently, even agents possessing identical information may legitimately arrive at different interpretations because their epistemic perspectives differ.

Knowledge claims produced through finite and perspectival cognition subsequently enter processes of

social mediation, where they are communicated, evaluated, criticised, and revised within epistemic communities. Scientific inquiry, legal reasoning, educational institutions, and digital communication platforms all illustrate that knowledge is rarely generated or validated in isolation. Rather, epistemic reliability emerges through collective practices of criticism, peer evaluation, institutional norms, and public accountability. At the same time, these social processes introduce their own vulnerabilities, including misinformation, algorithmic amplification, unequal credibility allocation, and institutional bias, thereby demonstrating that social validation cannot itself guarantee certainty.

Finally, every attempt to justify the standards governing knowledge evaluation encounters epistemic recursion. The criteria used to assess evidence, determine reliability, or establish justification themselves require justification, creating an unavoidable regress of epistemic evaluation. Rather than representing a defect of reasoning, this recursive structure reveals that epistemic inquiry is inherently self-corrective and permanently open to revision. Every epistemic framework must therefore acknowledge the provisional character of its own justificatory standards.

The conceptual framework is presented in Figure 1, which depicts the five stages of knowledge building, mutually connected. The impossibility of full information, the influence of perspectivity on information, the mediation of social information, and the recursive character of epistemic evaluation of the criteria established to justify claims of knowledge. These processes result in the updating of knowledge claims when new evidence, views, or approaches come to light. The feedback loops show knowledge as a cyclic activity, not a linear one. The human and AI examples illustrate the way that the same structural limitations also hold for all finite epistemic agents, biological and artificial, and therefore reinforce the argument in the paper that absolute epistemic certainty is impossible for all agents of finite information.

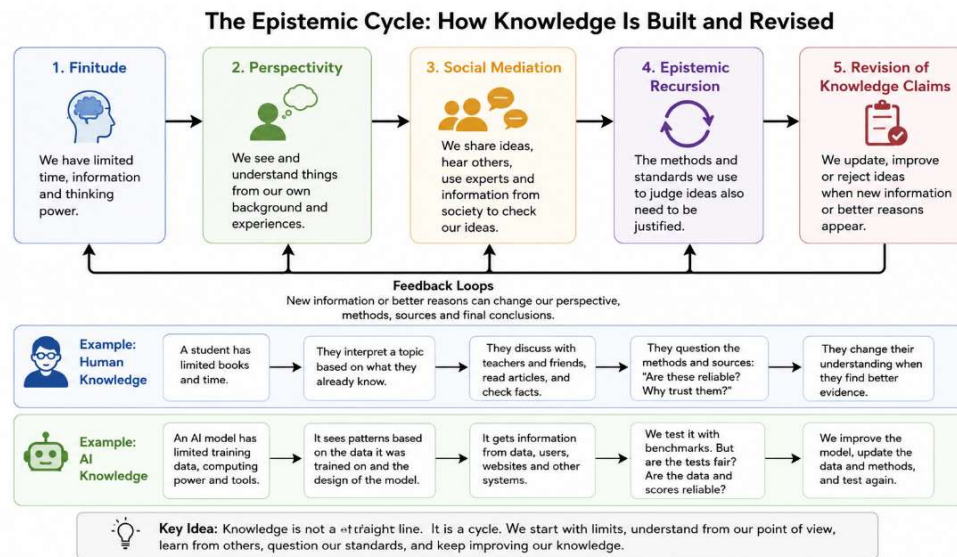


Figure 1: The Epistemic Cycle of Finite Informational Systems.

6. FALLIBILISM AND THE REGULATIVE IDEAL OF CERTAINTY

The integrated framework developed in the preceding section provides the philosophical basis for a fallibilist interpretation of knowledge. If finitude, perspectivity, social mediation, and epistemic recursion constitute unavoidable structural features of finite epistemic systems, then certainty cannot reasonably function as a prerequisite for knowledge. Instead, the appropriate epistemological response is fallibilism, which recognises that justified knowledge remains compatible with the continuing possibility of error.

6.1 The Fallibilist Position

Fallibilism holds that knowledge does not need absolute certainty, only the presence of sufficiently justified belief, even when there is the possibility of error (Popper, 2005); (Williamson, 2002). The present framework steps in by suggesting the idea of structural fallibilism, which states that error exists not only because of the incompleteness of evidence, but because all finite epistemic systems have the embedded properties of finitude, perspectivity, social mediation, and epistemic recursion. The reason why certainty is not achievable, for either human or artificial agents, is due to these structural constraints.

6.2 Certainty as a Regulative Ideal

Absolute certainty is never reached for a finite informational agent, so certainty must be considered a regulative ideal and not an epistemic state. In such a way, a regulative ideal serves as a norm, a normativity that informs the investigation process to ask for continual improvements, to reflect critically, and to be open to revision without any pretense of certainty. Inquiry is an approach that leads to increasingly reliable knowledge by continually evaluating, not final, infallible knowledge. (Paseri & Durante, 2025). In this way, the proposed framework situates certainty between scepticism and infallibilism: while it does assume that knowledge is possible, it does not require certainty as a condition of knowledge, as does infallibilism. This conception is based on structural fallibilism: finitude, perspectivity, social mediation, and epistemic recursion account for why there can never be a complete absence of the possibility of error. Epistemic responsibility, then, implies openness to revision, transparency with respect to evidence, and explicit justification of the methods. The same applies to artificial intelligence, where systems should communicate uncertainty, disclose methodological limitations, and allow for constant updates of outputs instead of claiming them to be infallible (Hauswald, 2025).

6.3 Against Infallibilism

Infallibilism, the doctrine that knowledge demands certainty, is the origin of theoretical scepticism: if the knowledge must exclude all possibility of error, then the possibility of error is enough for there not to be any knowledge. Infallibilism can give rise to typical distortions at the practical level. Daniel Kahneman's work on cognitive psychology has confirmed that motivated reasoning – the desire to find and understand information that reinforces existing beliefs – is a phenomenon well documented, especially in domains where stakes are high, and uncertainty threatens, such as the domain of gambling. (Kahneman, 2011). (Kirmayer, 2024) has suggested that the social nature of digital information environments exacerbates such tendencies toward motivated reasoning, establishing epistemic communities with shared certainty over the information they encounter, which are shielded from external opposition. When certainty is not possible, the need for it puts pressure to resort to a psychological alternative, motivated reasoning, in which the agent creates certainty that they are unable to rationally find.

6.4 Addressing Alternative Epistemological Perspectives

A proposed framework should be viewed as additional, not alternative, to known epistemological theories. Foundationalists believe that properly basic beliefs break the circular regress of justification and that they offer solid foundations for knowledge. Such beliefs may help stifle inquiry, but they are limited in interpretation and application by the limits of cognition and social context. Coherentists believe that justification is based on the mutual support of a coherent set of beliefs. The present framework acknowledges that coherence contributes to rational justification, but claims that coherence can be as strong yet conflicting as other coherent but incompatible systems can arise under other informational conditions. Reliabilists argue that knowledge requires reliable cognitive processes but does not require certainty. This insight is captured, but the framework asks what the processes of reliability are and how it is assessed and continually reviewed. Contextualists describe the difference in knowledge standards across contexts but don't infer from that that there is a relativising claim. Structural epistemic constraints can be lessened, but not fully removed, by intellectual virtues like

open-mindedness and intellectual humility, which are the emphasis of virtue epistemologists. In terms of epistemology, knowledge-first epistemology sees knowledge as the starting point, but doesn't explain much about the conditions of the structure of knowledge acquisition and revision. Lastly, hinge epistemology has held that commitments to inquiry are tacit and unquestionable; the conception of the present framework sees them more as practical starting points than as sources of absolute certainty. Together, these theories shed light on various facets of knowledge, and the four-limit framework offers a general account of continued structural conditions which give rise to epistemic certainty in human and artificial epistemic agents.

7. IMPLICATIONS FOR ARTIFICIAL EPISTEMIC AGENTS

Rather, this study does not view large language models (LLMs) as epistemic entities but rather as information-processing and decision-support systems, which are increasingly being used as epistemic authorities when used in high-stakes situations. This is significant because the LLM's beliefs, intentions, and understanding are not in the human sense, but rather probabilistic outputs that affect human knowledge and decision-making. Thus, their epistemic significance is not in knowing, but in human epistemic practices. It can be demonstrated using LLMs for legal advice. The finitude is observed since the outputs of the model depend on the finite training data, architectural restrictions, retrieval quality, context windows, retrieval procedures, and available legal resources. Therefore, the model might not cover some recent legislation or precedent from the jurisdiction, which might be a problem in terms of accuracy (Ji et al., 2023). Perspectivity is created by the legal bodies, which in turn are shaped by the legal traditions, language norms, and institutional practices of the community, and thus are not a neutral rendition of the law. Social mediation is evident because the model does not validate legal knowledge, but the courts, the professional model, the legislation, and the interpretation of the experts do. An LLM may then reflect or perpetuate socialized biases, or offer conflicting legal interpretations as fact (Bender et al., 2021). Lastly, epistemic recursion may be observed since the value of legal advice will rely on various benchmarks whose validity is dependent on human labeling decisions, scoring criteria, expert judgments, and assumptions about legal correctness, all of which must be independently justified (Paseri & Durante, 2025). This may be compared to human cognition, except that it is not equivalent. The difference between consciousness, intentionality, and reasoning between humans and AI is fundamental. Both, however, work with bounded information systems, whose output is limited by the information available, the local interpretation thereof, socially mediated validation, and changeable criteria for assessment. The proposed framework is therefore able to describe structurally similar but functionally analogous constraints, rather than equivalent cognitive processes, which is why the acquisition of absolute epistemic certainty is impossible for both humans and machines.

8. IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH

Here, the main significance of the proposed four-limit structure for present-day epistemology, AI, and the philosophy of information is explored. It details the theoretical impact of the study, addresses the conceptual constraints, and suggests other avenues for future research. The section's examination of the framework's philosophical argument is not limited to the specific context of AI, but also shows how it can contribute to current debates on knowledge, certainty, epistemic responsibility, and the role of AI in decision-making, and how it may be empirically tested and developed by interdisciplinary research.

8.1 Theoretical Contributions

The principal theoretical contribution of this study lies in the development of an integrated philosophical framework that reconceptualises epistemic certainty as a structural rather than merely psychological or methodological problem. While previous scholarship has examined certainty through foundationalism, scepticism, fallibilism, information theory, or artificial intelligence independently, this study demonstrates that these traditions converge upon four interrelated structural limits

governing every finite informational agent: finitude, perspectivity, social mediation, and epistemic recursion. By integrating these dimensions within a single explanatory framework, the paper extends contemporary fallibilist epistemology beyond discussions of justification alone and provides a broader account of the informational architecture within which knowledge is acquired, evaluated, and revised. Consequently, the proposed framework offers a unified philosophical perspective capable of explaining epistemic limitations across both biological cognition and artificial intelligence.

8.2 Practical Implications

The suggested framework provides for four practical implications. Uncertainty disclosure and verification of knowledge claims using tools are required for finitude. Perspectivity emphasizes the importance of multiple data sources, incorporating multiple perspectives and the involvement of stakeholders in the evaluation of knowledge. Social mediation focuses on traceability, open sources, and meaningful supervision by people. Epistemic recursion necessitates constant monitoring of the audits, the regulation of the evaluation process, and a periodic reconsideration of the evaluation process and validation methods, so that both human and AI knowledge systems can be relied upon to be reliable, transparent, and subject to revision.

8.3 Limitations of the Study

Several limitations should be acknowledged. First, the present investigation is a conceptual and philosophical inquiry rather than an empirical study; accordingly, the proposed framework has not been empirically tested across specific informational environments or artificial intelligence systems. Second, the framework intentionally emphasises broad structural characteristics common to finite epistemic agents and therefore does not examine domain-specific variations that may arise within scientific, legal, medical, or educational contexts. Third, although the discussion incorporates contemporary developments in artificial intelligence, the rapid evolution of AI technologies may require future refinement of particular applications while leaving the broader philosophical framework intact. The proposed framework is conceptual, has not been empirically validated, and may require refinement when applied across different epistemic domains and disciplinary contexts.

8.4 Future Research Directions

Future research should investigate the explanatory and practical value of the proposed framework across diverse epistemic domains. Empirical studies could examine how the four structural limits influence decision-making, scientific collaboration, institutional knowledge production, and human interaction with increasingly sophisticated AI systems. Comparative philosophical research may further explore the applicability of the framework to non-Western epistemological traditions, thereby extending its conceptual scope beyond the predominantly Western perspectives considered in this study. Additionally, interdisciplinary collaboration among philosophers, information scientists, computer scientists, cognitive psychologists, and AI ethicists may contribute to developing operational indicators capable of evaluating epistemic responsibility within both human and artificial informational systems. Such investigations would not only strengthen the theoretical foundations established here but also support the responsible development of future knowledge technologies in increasingly complex digital societies.

9. CONCLUSION

This paper has suggested that none of these finite informational agents can claim absolute epistemic certainty, because knowledge is bounded by four interrelated structural limitations, namely: finitude, perspectivity, social mediation, and epistemic recursion. Nonetheless, the study draws from, and is informed by, classical epistemology, philosophy of information and social epistemology, as well as more recent knowledge from the field of AI, to show that these limitations are not incidental features of limited informational systems, but essential aspects of them. Certainty should, then, be regarded as a regulative ideal, one that must be continually sought, critically examined, and revised responsibly.

This view engenders a fallibilist approach to knowledge that is based on structure and seeks to allow for the acquisition of true knowledge without claiming that it is impossible to eliminate epistemic error entirely. The framework also has important implications for artificial intelligence. The study does not view AI systems as independent knowers, but rather as information-processing and decision-support systems that need to be communicated with openly, monitored and supervised by humans, and continually evaluated. Epistemically responsible design and governance are thus essential for responsible AI as well as for the technical aspects of AI performance. This paper, in theory, offers a solution to the unresolved issue of why different epistemological inadequacies have resulted from the same structural inadequacies in both human and artificial information systems. The suggested four-limit structure provides a common explanatory model that combines classical epistemology, philosophy of information, and AI epistemology beyond any domain-specific models. This is important because they offer a principled basis for designing responsible AI-mediated knowledge practices that are transparent, accountable, revisionary, and epistemically responsible, not presuming certainty.

REFERENCES

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?? Proceedings of the 2021 ACM conference on fairness, accountability, and transparency,
- Cabrera, F. (2023). Inference to the best explanation: An overview. *Handbook of abductive cognition*, 1863-1896.
- Calvente, S. (2022). David Hume on the True Philosopher's Identity. *Diálogos*.
- Carroll, T. D. (2025). Rethinking Philosophy of Religion with Wittgenstein: Religious Diversities and Racism.
- Chesney, B., & Citron, D. (2019). Deep fakes. *California law review*, 107(6), 1753-1820.
- Feldman, J. (2024). Gañesa and the Gettier problem. In *The Vindication of the World* (pp. 117-134). Routledge.
- Floridi, L. (2009). Philosophical conceptions of information. In *Formal theories of information: From Shannon to semantic information theory and general concepts of information* (pp. 13-53). Springer.
- Floridi, L. (2013). *The philosophy of information*. OUP Oxford.
- Goldman, A. (2025). Reliabilism. *A companion to epistemology*, 1, 367-375.
- Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese*, 201(3), 103.
- Harding, S. G. (1986). *The science question in feminism*. Cornell university press.
- Hauswald, R. (2025). Artificial epistemic authorities. *Social Epistemology*, 39(6), 716-725.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM computing surveys*, 55(12), 1-38.
- Kahneman, D. (2011). Thinking, fast and slow. *Farrar, Straus and Giroux*.
- Kaya, A., & Benevento, A. (2022). Epistemic justice as a political capability of radicalised youth in Europe: a case of knowledge production with local researchers. *Journal of Human Development and Capabilities*, 23(1), 73-94.
- Kirmayer, L. J. (2024). Science and sanity: A social epistemology of misinformation, disinformation, and the limits of knowledge. *Transcultural psychiatry*, 61(5), 795-808.
- Kuhn, T. S. (1997). *The structure of scientific revolutions* (Vol. 962). University of Chicago press Chicago.
- Landes, J. (2022). Bayesian Epistemology. *KRITERION—Journal of Philosophy*, 36(1), 1-7.

- Levy, N. (2021). Bad beliefs: Why they happen to good people.
- Naser, M. (2025). A guide to machine learning epistemic Ignorance, hidden Paradoxes, and other tensions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 15(3), e70038.
- Paseri, L., & Durante, M. (2025). Examining epistemological challenges of large language models in law. *Cambridge Forum on AI: Law and Governance*,
- Paul, E. S. (2024). The Rational Force of Clarity: Descartes's Rejection of Psychologism.
- Paulson, S. (2024). *Epistemic Reasons & Cognitive Self-Monitoring* Northwestern University].
- Popper, K. (2005). *The logic of scientific discovery*. Routledge.
- Rowland, F. (2011). The filter bubble: What the internet is hiding from you. *portal: Libraries and the Academy*, 11(4), 1009-1011.
- Sosa, E. (2021). *Epistemic explanations: A theory of telic normativity, and what it explains*. Oxford University Press.
- Stein, E. Simon, HA (1957) *Models of Man: Social and Rational*. New York: Wiley. Simon, HA (1983) *Reason in Human Affairs*. Stanford, CA: Stanford University Press. Sloman, SA (1996) The empirical case for two systems of reasoning, *Psychological Bulletin*, 119, 3-22. Sloman, SA & Over, DE (2003) Probability judgment from the inside and out. In DE Over. *Science*, 16, 1-40.
- Sunstein, C. R. (2018). *Republic: Divided democracy in the age of social media*.
- Titelbaum, M. G. (2022). *Fundamentals of Bayesian epistemology 1: Introducing credences*. Oxford University Press.
- Williamson, T. (2002). *Knowledge and its Limits*. OUP UK.
- Spivack, N. (2025, June 11). Epistemology and metacognition in artificial intelligence: Defining, classifying, and governing the limits of AI knowledge. *Nova Spivack Blog*. <https://www.novaspivack.com/technology/ai-technology/epistemology-and-metacognition-in-artificial-intelligence-defining-classifying-and-governing-the-limits-of-ai-knowledge>