A DIGITAL SPACE LAB Publication

Epistemic Boundaries of Large Language Models

Explanation, Orientation, and Structural Limits

Harald Meier
Independent Researcher · Digital Space Lab

Version 1.1 · 2026

Abstract

Large language models (LLMs) have demonstrated a remarkable ability to generate coherent, context-sensitive, and seemingly reasoned outputs. This has led to their increasing use in contexts that extend beyond information retrieval toward decision support. This paper argues that such uses rest on a fundamental epistemic misunderstanding. It introduces a distinction between explanation and orientation to clarify the structural limits of LLMs. While LLMs expand the space of possible representations by differentiating elements and articulating relations, they do not perform the integration required to stabilize relevance as a condition for action. As a result, improvements in coherence and reasoning transparency do not constitute progress toward decision-making capacity, but increase the risk that explanation is mistaken for orientation. This misinterpretation shifts epistemic risk from system failure to system success. By situating LLMs within the domain of explanation, the paper establishes an epistemic boundary that redefines the conditions under which judgment and responsibility can meaningfully operate beyond epistemic processing.

Keywords

Large Language Models · Epistemic Boundaries · Explanation · Orientation · Human Judgment · Decision-Making

CONTENTS

1. Introduction

Large language models (LLMs) have rapidly transformed the landscape of artificial intelligence by demonstrating an unprecedented ability to generate coherent, context-sensitive, and structurally complex outputs. Across domains, these systems produce responses that resemble human reasoning in both form and depth, leading to their increasing use in contexts that extend beyond information retrieval into areas approaching judgment and decision-making. As a result, their outputs are not only consumed as explanations but are often implicitly treated as if they were sufficient for guiding action.

This shift introduces a fundamental epistemic problem that cannot be reduced to questions of accuracy, bias, or hallucination. While these concerns remain central to the evaluation of model performance, they operate within the domain of output quality and do not address the structural nature of what is being produced. The critical issue is therefore not located at the level of output quality, but at the level of interpretation, where explanatory outputs are treated as if they already satisfy the conditions required for decision-making. As a result, the central epistemic risk does not arise from system failure, but from system success. The more coherent and persuasive these outputs become, the more likely they are to be misinterpreted as providing the conditions required for action.

This paper argues that this interpretation is structurally misleading. LLMs operate within the domain of explanation, understood as the differentiation of elements and the articulation of relations, but they do not perform orientation, defined here as the integration of constraints into a structured configuration that stabilizes relevance under conditions of tension. Decision, in turn, introduces commitment under conditions of irreversibility and cannot be derived from either explanation or orientation, even though it presupposes both. The distinction between these domains

is not merely conceptual but structural, as each corresponds to a different mode of epistemic processing that cannot be reduced to the others without distortion.

The relevance of this distinction becomes particularly visible in the context of contemporary AI discourse. Much of the current debate assumes that increasingly sophisticated forms of reasoning, such as those enabled by chain-of-thought prompting (Wei et al., 2022), bring artificial systems closer to decision-making capacities. However, these approaches remain confined to the expansion of explanation, as they generate more detailed inferential sequences without integrating constraints into a binding configuration. Similarly, critiques of large language models, such as the characterization of these systems as "stochastic parrots" (Bender et al., 2021), highlight important limitations but do not fully account for the structural distinction developed here.

To clarify this distinction, the argument draws on philosophical and systems-theoretical perspectives while extending them beyond their original scope. In Kantian terms, one can distinguish between the generation of representations and the conditions under which these representations become binding for action (Kant, 1781/1998). In system-theoretical terms, a similar distinction can be drawn between the expansion of possible observations and the reduction of this space through the integration of constraints into a structured configuration (Luhmann, 1995). These perspectives provide important points of reference but do not fully capture the operational distinction introduced here, which isolates orientation as an independent epistemic operation.

The contribution of this paper is twofold. First, it introduces a precise distinction between explanation, orientation, and decision, thereby clarifying a structural dimension that remains underdeveloped in current AI discourse. Second, it applies this distinction to large language models, demonstrating why their capabilities remain confined to explanation and why this limitation cannot be overcome through incremental improvements in reasoning performance. This reframing shifts the focus from what these systems can do to how their outputs are interpreted and what epistemic operations are attributed to them.

The argument proceeds as follows. Section 2 introduces the epistemic architecture by defining the three epistemic domains. Section 3 specifies the internal structure of orientation. Section 4 analyzes large language models as systems of explanation and examines the limits of their capabilities. Section 5 identifies the epistemic risks that arise from misinterpreting explanation as orientation. Section 6 discusses the implications for AI design and use, while Section 7 concludes by emphasizing that the central challenge is not the insufficiency of these systems, but the conditions under which they are understood.

2. Epistemic Architecture: Explanation, Orientation, and Decision

The argument developed in this paper rests on a distinction between three irreducible epistemic domains: explanation, orientation, and decision. These domains do not represent successive stages in a linear process, nor can they be reduced to one another. Rather, they constitute structurally distinct forms of epistemic operation that fulfill different functions and operate under different conditions. Understanding this distinction requires a shift in perspective from representation to operation, that is, from what is described to how epistemic processing is structured. By operation, this paper refers to structurally distinct ways in which relevance is produced and constrained, rather than to cognitive or procedural steps. In this sense, the framework aligns with system-theoretical approaches that emphasize the primacy of operations over contents (Luhmann, 1995), while extending them by introducing a more precise account of how binding relevance emerges prior to decision. Related concerns appear in discussions of judgment, situated interpretation, and sociotechnical decision-making, though the distinction developed here isolates orientation as a distinct epistemic operation.

Explanation refers to the differentiation of elements and the articulation of relations between them. It produces structured representations by identifying relevant factors, connecting them through causal, logical, or associative relations, and organizing them into coherent descriptions or inferential sequences. In the context of large language models, this capacity is particularly visible, as these systems generate outputs that connect premises, evidence, and conclusions in ways that are syntactically well-formed and semantically meaningful. Techniques such as chain-of-thought prompting further amplify this capacity by making intermediate reasoning steps explicit, thereby enhancing the transparency of the explanatory process (Wei et al., 2022). However, explanation does not establish which elements are binding under given conditions. It differentiates and relates, but it does not determine relevance in a way that constrains action. Multiple explanatory structures can coexist without resolving which of them should guide decision-making. Explanation expands the space of possible understanding without reducing it to a configuration that constrains what can count as relevant.

Orientation operates on a different epistemic condition. It becomes necessary where explanation leaves relevance underdetermined and where no extension of explanatory differentiation resolves this indeterminacy. Orientation does not extend explanation, nor does it arise through gradual accumulation of information, but through a discontinuous change in epistemic operation. It integrates relations under constraints in a way that stabilizes relevance. Constraints do not merely restrict a pre-given space of possibilities but act as structuring forces that shape how elements become relevant to one another. Because constraints are often heterogeneous and incompatible, orientation necessarily involves the organization of tension rather than

its elimination. It is this integration of constraints, rather than the accumulation of relations, that enables orientation. Orientation produces binding without commitment: it establishes a structured field of relevance without collapsing the plurality of possibilities into a single outcome.

Decision, by contrast, introduces commitment under conditions of irreversibility. While explanation and orientation remain within the domain of epistemic structuring, decision transforms the situation by selecting one course of action from a set of possibilities and thereby closing the space of alternatives. This closure cannot be derived from explanation, nor is it determined by orientation, even though orientation provides the conditions under which decision becomes meaningful. Decision therefore marks a discontinuity that cannot be bridged through increased informational richness or structural complexity alone. In doing so, it introduces responsibility that cannot be resolved within epistemic structuring.

The structural relations between these domains can be expressed more precisely: explanation produces differentiation without binding, orientation produces binding without commitment, and decision introduces commitment under conditions of irreversibility. These domains are irreducible because they fulfill different epistemic functions and operate under different constraints. Attempts to collapse them into one another lead to systematic distortions, particularly when explanation is treated as sufficient for decision or when orientation is assumed to determine commitment. Such collapses obscure the conditions under which relevance is stabilized and responsibility is assumed.

The transition from orientation to decision is defined here as termination. Termination does not denote a gradual threshold but a boundary at which a configuration of relevance becomes sufficient to support commitment under conditions of irreversibility. It is not a transition within epistemic processing, but the point at which epistemic operations cease to determine the situation. At this boundary, epistemic structuring gives way to responsibility, as commitment introduces consequences that cannot be resolved within epistemic operations alone. The boundary remains discontinuous: no degree of explanatory richness or orientational complexity can, by itself, produce the act of commitment. Stabilization creates the condition under which termination may occur, but it does not determine it.

Orientation, in this sense, is not a cognitive faculty or a psychological process, but a structural operation that can be analytically reconstructed across different domains of epistemic processing. It does not depend on a specific agent but on the configuration of relations under constraint that stabilizes relevance prior to decision. The presence or absence of orientation can therefore be described in structural terms rather than inferred from the apparent coherence of outputs.

While this establishes orientation as a distinct epistemic operation, its internal structure remains to be specified. The following section develops this structure in greater detail (see Figure 1 for an overview of the epistemic architecture).

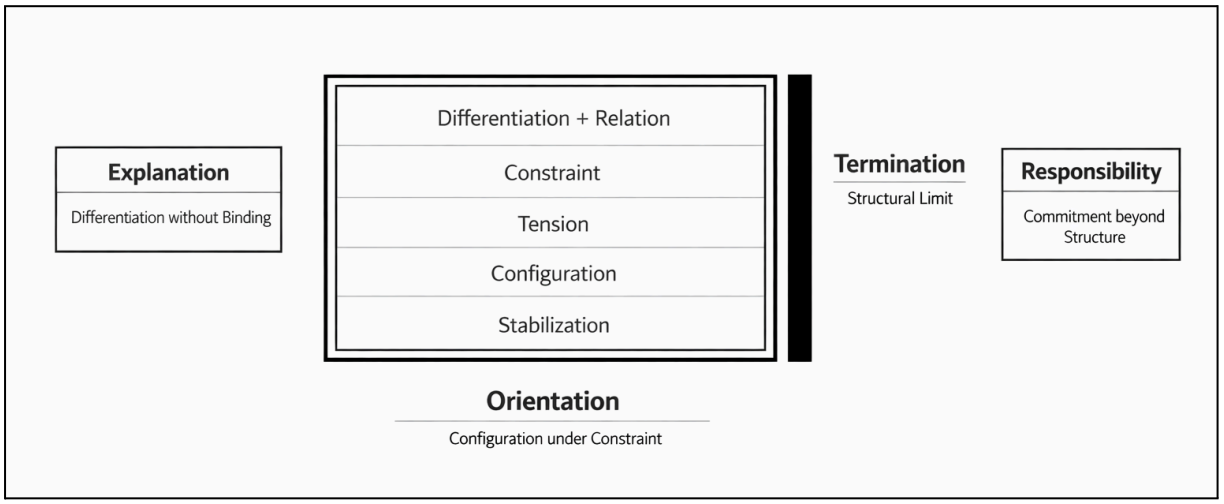

Figure 1. Epistemic architecture of explanation, orientation, and decision.

3. Internal Structure of Orientation

Having introduced orientation in functional terms, its internal structure can now be specified more precisely. Orientation is not a discrete act, nor a procedural step that follows from explanation, but a configuration that emerges under conditions of constraint. In this configuration, differentiated elements, constraints, and tensions are held together in a way that stabilizes relevance. Where explanation produces relations without binding, orientation introduces a form of integration that renders certain relations structurally decisive without collapsing the field into immediate action. Integration, in this sense, does not denote a generic synthesis, but the binding of relations under constraint such that they become structurally decisive.

This configuration cannot be understood as a sequence of steps. Instead, orientation is composed of interdependent structural dimensions that are co-present and mutually conditioning. Differentiation and relation provide the material inherited from explanation: elements are distinguished and connected, yet not yet integrated into a binding structure. Tension emerges as the central organizing principle wherever constraints cannot be simultaneously satisfied and must therefore be held in relation. It is not a disturbance to be eliminated, but a constitutive feature that organizes the field and necessitates integration. Configuration refers to the structuring through which elements are bound together under constraint; it is not a descriptive arrangement, but a binding structure that renders relations operative

within a given situation. Stabilization marks the condition in which tensions are not resolved but sufficiently integrated such that the configuration becomes structurally coherent and capable of supporting commitment without determining it.

Orientation is therefore not a static state, but a structured dynamic in which stabilization remains provisional and subject to reconfiguration. Stabilization does not arise from the accumulation of information, nor from the extension of inferential chains, but from the ongoing reconfiguration of tensions under constraint. This dynamic remains inherently open-ended: configurations may shift, tensions may reappear in new forms, and stabilization remains provisional. What distinguishes orientation is not closure, but the temporary achievement of coherence under conditions that cannot be fully harmonized.

Different structural conditions can be distinguished without thereby introducing a procedural method of assessment. The presence or absence of orientation can be described in structural terms rather than inferred from output quality alone. Situations may remain underconstrained, where relevance is not stabilized; overconstrained, where incompatible constraints prevent integration; or unstable, where provisional configurations fail to sustain coherence. Only in the case of stabilization does orientation occur, when tensions are held together in a coherent configuration that renders certain relations decisively relevant. These distinctions are not criteria to be mechanically applied, but indicators of the structural condition of the configuration itself.

The relation between orientation and decision must be drawn with equal care. Stabilization is a necessary condition for decision, but it is not sufficient. Stabilization creates the condition under which termination may occur, but does not determine it. Termination arises only when a configuration reaches a level of sufficiency that allows for commitment under conditions of irreversibility. No configuration, regardless of its stability, entails termination. At this boundary, epistemic structuring gives way to responsibility, as the consequences of commitment extend beyond what can be resolved within orientation. The boundary between orientation and decision therefore remains discontinuous.

This architecture will serve as a framework for understanding the epistemic limits of large language models in the following section.

4. Large Language Models and the Limits of Explanation

Large language models (LLMs) have demonstrated a remarkable capacity to generate coherent, context-sensitive, and structurally complex outputs across a wide range of domains. Their ability to produce explanations that resemble human reasoning in both form and depth has led to their increasing use in contexts that extend beyond information retrieval into areas approaching judgment and decision-making. As these

systems become more integrated into professional and societal contexts, their outputs are not only interpreted as explanations but are increasingly treated as if they were sufficient for guiding action, thereby raising the question of what kind of epistemic operation they actually perform.

4.1 LLMs as Systems of Explanation

From a structural perspective, LLMs operate as systems of explanation. They generate outputs by predicting the most probable continuation of a sequence given a context, drawing on statistical regularities learned from large-scale training data (Bender et al., 2021; Bommasani et al., 2021). This process enables them to produce responses that are syntactically well-formed, semantically coherent, and often rhetorically persuasive, creating the impression of structured reasoning even in complex domains.

Recent developments have further enhanced this explanatory capacity. Techniques such as chain-of-thought prompting allow models to generate multi-step reasoning processes that approximate structured argumentation by making intermediate steps explicit (Wei et al., 2022), while approaches such as self-consistency decoding improve the reliability of these reasoning chains by aggregating multiple sampled outputs (Wang et al., 2022). These methods significantly expand the domain of explanation, enabling LLMs to differentiate elements, articulate relations, and construct inferential sequences with increasing sophistication.

However, these advances remain confined to the domain of explanation. They improve the structure and transparency of generated outputs but do not transform these outputs into configurations that integrate constraints, organize tensions, or stabilize relevance in a binding sense. What is enhanced is the clarity of representation, not the structure of relevance on which action depends.

4.2 The Illusion of Orientation

The increasing coherence of LLM outputs creates a structural risk: explanation is misinterpreted as orientation. When a response is internally consistent, contextually appropriate, and rhetorically persuasive, it produces the impression that the epistemic work required for decision-making has already been performed. This impression is reinforced by the ability of LLMs to incorporate multiple perspectives, articulate trade-offs, and simulate deliberative reasoning.

Yet this resemblance is misleading. Constraints appear in LLM outputs as informational elements rather than as structuring conditions that limit possible configurations. Tensions are described but not organized in a way that establishes stable relations of priority or dominance. The result is a structured representation that remains open, offering multiple plausible pathways without establishing which of

them should guide action. The output may appear complete, but the underlying field of relevance remains unstructured.

This misinterpretation is not incidental but structurally inevitable. It does not imply that LLMs cannot simulate aspects associated with orientation, such as the articulation of constraints or the representation of trade-offs. However, such simulations remain confined to the domain of explanation, as they do not produce binding configurations of relevance under constraint. The appearance of orientation is therefore generated within explanation rather than through a transformation beyond it.

This raises the question of whether more complex system architectures might overcome this limitation. A potential objection concerns the emergence of so-called agentic systems that combine large language models with external tools, feedback loops, and environmental interaction. Such systems incorporate constraints through external state tracking, iterative refinement, and goal-directed execution. However, this does not alter the structural limitation identified here. Constraints remain representational inputs that guide subsequent outputs, rather than structuring conditions that stabilize relevance across a field of possibilities. Even when extended through tool use or iterative loops, these systems operate through the generation of plausible continuations rather than the formation of binding configurations. What changes is the complexity of representation, not the structure of relevance on which orientation depends.

The inability of LLMs to perform orientation follows directly from this condition. Orientation requires the integration of constraints as binding conditions, the organization of tensions into stable configurations, and the stabilization of relevance in a way that reduces the space of possibilities without eliminating it. These operations involve a form of epistemic closure that is incompatible with the open-ended, reversible, and generative nature of probabilistic sequence modeling.

From a philosophical perspective, this limitation can be understood as a boundary of representation. In Kantian terms, LLMs operate within the domain of possible representations but lack the capacity to determine the conditions under which these representations become binding for action (Kant, 1781/1998). From a systems-theoretical perspective, they expand the space of possible observations without performing the reduction of this space through binding selections that establish relevance (Luhmann, 1995). Both perspectives converge on the insight that the expansion of representational capacity does not entail the capacity for orientation.

Situating this distinction in relation to existing approaches further clarifies its significance. Chain-of-thought prompting enhances the transparency of reasoning but remains confined to explanation, as it makes intermediate steps visible without integrating them into a binding configuration. Bounded rationality (Simon, 1957)

addresses decision-making under constraint but presupposes a structured field of relevance rather than explaining its formation. Even approaches such as human-compatible AI (Russell, 2019), which aim to align systems with human values, operate at the level of output optimization and do not distinguish between explanatory generation and orientational structuring. Taken together, these approaches reveal a consistent pattern: they address problems that implicitly require orientation while remaining grounded in explanation.

5. Epistemic Risks

The distinction between explanation and orientation reveals a form of epistemic risk that cannot be adequately captured by existing discussions of accuracy, bias, or hallucination. While these issues remain important, they are located within the domain of explanation and therefore do not address the structural problem that arises when explanatory outputs are interpreted as if they already provide orientation. The central risk is not that large language models produce incorrect outputs, but that their outputs are taken to be more than they are. The risk emerges not from error, but from epistemic over-attribution.

This misinterpretation introduces a structural instability into decision-making processes. When explanatory coherence is treated as a substitute for orientational sufficiency, the conditions under which action is taken remain insufficiently structured. Constraints may be mentioned but are not integrated as binding conditions, tensions may be articulated but are not organized into stable configurations, and relevance may appear to be established without being structurally stabilized. As a result, decisions may be made on the basis of configurations that remain epistemically open, even though they appear complete.

This structural instability becomes particularly visible in high-stakes contexts. For example, in a medical decision scenario, an LLM may generate a coherent explanation that outlines treatment options, associated risks, and probabilistic outcomes. While this output can be highly informative, it does not integrate patient-specific constraints, contextual priorities, or institutional limitations into a binding configuration of relevance. The result is a structured representation that appears decision-ready but remains epistemically open, requiring orientation that the system itself does not perform.

The increasing coherence of LLM outputs amplifies this risk. As these systems improve, their responses become more consistent, contextually aligned, and rhetorically persuasive, making it increasingly difficult to distinguish between explanatory richness and orientational structure. The problem is not complexity but misinterpretation, as the apparent completeness of the output masks the absence of

binding integration. In this sense, the locus of risk shifts from system failure to system success.

This shift has direct implications for how decisions are formed. When the distinction between explanation and orientation collapses, decisions may occur without the prior stabilization of relevance that orientation provides. Such decisions are not necessarily incorrect in a factual sense, but they are structurally misaligned with the conditions under which they are made. The absence of orientation does not produce immediate error, but increases the likelihood of misalignment over time, particularly in complex environments where constraints and tensions cannot be reduced to simple trade-offs.

At the same time, this misinterpretation affects the distribution of responsibility. If explanatory outputs are treated as if they already incorporate orientation, the boundary between epistemic processing and decision-making becomes blurred. Responsibility may then be implicitly displaced from human actors to the system, even though the system does not perform the operations required for decision and thus cannot cross the termination boundary. This displacement is not the result of deliberate delegation, but of a structural misunderstanding of what the system provides.

In this context, the concept of judgment becomes critical. Judgment involves the capacity to act under conditions of plurality and uncertainty, integrating constraints and assuming responsibility for consequences. As Arendt (1958) emphasizes, judgment cannot be reduced to rule-following or procedural reasoning, but requires the ability to relate particular situations to a broader field of meaning while remaining accountable for the outcome. Systems that generate explanations without performing orientation cannot assume this role, and treating them as if they could risks undermining the conditions under which responsibility can be meaningfully exercised.

Because decisions introduce commitment under conditions of irreversibility, any distortion in the epistemic conditions under which they are formed necessarily extends into the ethical and political domain. The epistemic risk identified here is therefore not limited to technical failure but concerns the conditions under which decisions are made, the structures that support or distort these conditions, and the ways in which responsibility is maintained or displaced. In this sense, the problem is not only epistemic but also ethical and political, as it affects how agency is distributed in systems where human and artificial processes are increasingly intertwined.

6. Implications for AI Design and Use

The distinction between explanation, orientation, and decision translates directly into implications for how large language models are designed, deployed, and interpreted. If LLMs are structurally limited to the domain of explanation, their integration into decision-making contexts must be guided by a clear understanding of what they can and cannot provide. This requires a conceptual reorientation that recognizes the epistemic boundary identified in this paper.

First, systems must be designed and evaluated in a way that reflects their explanatory nature. Current evaluation frameworks often treat coherence, plausibility, and correctness as indicators of decision-support capability, thereby obscuring the distinction between explanation and orientation. While such metrics improve the quality of outputs within the explanatory domain, they do not indicate whether these outputs are sufficient for guiding action. Evaluation practices must therefore avoid interpreting explanatory performance as a proxy for orientational sufficiency. This requires the explicit separation of explanatory generation and orientational structuring, for example by combining LLM-based outputs with processes that integrate constraints and organize tensions at the level of human or institutional decision-making.

Second, the interface between human and system must preserve the distinction between explanation and orientation. Rather than presenting outputs as if they were already structured for decision-making, systems should make explicit that they operate within the domain of explanation. This does not imply reducing their utility, but clarifying their epistemic role. A productive direction is the development of interfaces that foreground the openness of the explanatory field, highlighting alternative configurations and explicitly marking the absence of binding integration. Such interfaces would not attempt to simulate orientation, but would support human actors in performing it.

Third, the role of human judgment must be reinforced rather than replaced. If orientation is a distinct epistemic operation that cannot be performed by LLMs, then the responsibility for integrating constraints, organizing tensions, and stabilizing relevance remains with human actors. The use of LLMs in decision-making contexts should therefore be framed not as a delegation of judgment, but as an augmentation of explanatory capacity. Systems can expand the space of possible representations, but they cannot determine which of these representations should guide action.

Fourth, governance frameworks must account for the epistemic boundary identified here. Current discussions of AI governance often focus on issues such as bias, transparency, and accountability, which are essential but insufficient. Without a clear distinction between explanation and orientation, governance mechanisms risk addressing symptoms rather than structural causes. Policies that regulate the use of

LLMs in high-stakes contexts must therefore consider not only the quality of outputs, but also the conditions under which these outputs are interpreted and used for decision-making.

Finally, the design of AI systems benefits from explicitly incorporating the distinction between epistemic domains. Rather than attempting to extend LLMs toward decision-making capabilities through increasingly complex reasoning techniques, it may be more productive to develop complementary systems or processes that support orientation as a separate operation. This could involve hybrid architectures in which explanatory generation is combined with structured human or institutional processes that perform the integration of constraints and the organization of tensions. Such approaches do not eliminate the epistemic boundary, but operate within it by combining different forms of epistemic processing.

Taken together, these implications reinforce that the responsible use of large language models depends not on extending their capabilities, but on recognizing and maintaining the epistemic boundary that defines them. Without this clarification, systems risk being used in ways that exceed their structural capacities, leading to misaligned decisions and the erosion of responsibility.

7. Conclusion

This paper has argued that large language models are structurally confined to the domain of explanation and therefore cannot perform the epistemic operation of orientation required for decision-making. By distinguishing between explanation, orientation, and decision, it has been shown that increasing coherence and sophistication in LLM outputs do not overcome this limitation, but instead create the conditions for a systematic misinterpretation in which explanation is treated as if it were sufficient for action.

The central contribution of this paper lies in the identification of an epistemic boundary. This boundary does not arise from a lack of technical capability, but from the structural conditions under which LLMs operate. As probabilistic systems that generate plausible continuations, they expand the space of possible representations without performing the binding integration required to stabilize relevance under constraint. Orientation, by contrast, requires a form of epistemic closure that cannot be achieved through generative continuation alone.

Recognizing this boundary clarifies why approaches that focus on improving reasoning capabilities do not address the core limitation, as they remain within the domain of explanation. It also explains why increasing model performance amplifies epistemic risk, as the plausibility of outputs makes them more likely to be misinterpreted as providing orientation. The problem is therefore not one of

insufficient performance, but of structural misalignment between what the system produces and how it is used.

At the same time, this framework provides a basis for rethinking the role of LLMs in complex environments. Rather than treating them as systems that approximate human judgment, they can be understood as tools that enhance explanatory capacity while leaving the operations of orientation and decision intact. This perspective preserves the role of human judgment and responsibility, not as a limitation to be overcome, but as a necessary component of epistemic processing in contexts where action has consequences.

The broader significance of this argument extends beyond large language models. It highlights the importance of distinguishing between different epistemic operations and recognizing the boundaries that separate them. In environments characterized by increasing complexity and the proliferation of information, the ability to differentiate between explanation, orientation, and decision becomes critical. Without this differentiation, the expansion of explanatory capacity may be mistaken for the resolution of uncertainty, leading to decisions that are not adequately grounded in structured relevance.

Future research may build on this framework by exploring how orientation can be supported in practice, particularly in hybrid human–AI systems, and by examining how the distinction between epistemic domains applies to other forms of artificial intelligence. Such work would not aim to eliminate the epistemic boundary identified here, but to better understand its implications and to develop approaches that operate within its constraints.

References

Arendt, H. (1958). The Human Condition. Chicago: University of Chicago Press.

Bender, E.M., Gebru, T., McMillan-Major, A. and 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 (FAccT), pp. 610–623.

Bommasani, R. et al. (2021). On the Opportunities and Risks of Foundation Models. arXiv preprint arXiv:2108.07258.

Kant, I. (1781/1998). Critique of Pure Reason. Cambridge: Cambridge University Press.

Luhmann, N. (1995). Social Systems. Stanford: Stanford University Press.

Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. New York: Viking.

Simon, H.A. (1957). Models of Man: Social and Rational. New York: Wiley.

Wang, X. et al. (2022). Self-Consistency Improves Chain of Thought Reasoning in Language Models. arXiv preprint arXiv:2203.11171.

Wei, J. et al. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv preprint arXiv:2201.11903.

PUBLICATION RECORD

Title
Epistemic Boundaries of Large Language Models. Explanation, Orientation, and Structural Limits

Author
Harald Meier

Affiliation
Independent Researcher · Digital Space Lab

Location
Winterberg, Germany

Date
2026-04-12

Version
1.1 · 2026

Type
Analytical — Application of Epistemic Architecture

Position
Analytical Extension — Applies the Epistheon framework without extending its architecture

Scope
Analyzes the epistemic limitations of large language models by distinguishing between explanation, orientation, termination, and responsibility as structurally irreducible domains

Delimitation
Does not modify or extend the Epistheon architecture
Does not propose technical improvements to language models
Does not prescribe decisions or operational procedures

Relation to Epistheon
This work applies the Epistheon architecture but is not part of the Epistheon corpus

Repository
Digital Space Lab — Hub
https://digitalspacelab.com/hub

DIGITAL SPACE LAB

An independent research and publication framework
for orientation, structural clarity,
and responsibility in complex knowledge environments.

digitalspacelab.com

The Digital Space Lab operates independently.
If this work has been useful, informal support is welcome:

ko-fi.com/digitalspacelab