

**FROM RISK ASSESSMENTS TO KNOWLEDGE STATEMENTS:
EPISTEMIC UNCERTAINTY IN LLM-BASED RISK MANAGEMENT
FOR CRITICAL INFRASTRUCTURES**

Sebastian Grüneberg

University of Library Studies and Information Technologies
<https://doi.org/10.70300/DOFP5009>

Abstract: *The growing use of large language models (LLMs) in risk management is often justified by improved efficiency and better handling of unstructured data. However, the epistemic quality of the resulting risk statements remains poorly explored. This article investigates epistemic uncertainty in LLM-based risk management from a conceptual perspective, drawing on risk theory and epistemological concepts. LLM-generated risk statements are analyzed as knowledge statements and assessed in terms of their epistemic viability and knowledge strength through a theory-driven literature review.*

The analysis shows that LLMs do not reduce epistemic uncertainty; instead, they alter its visibility and assessability. Fluent risk statements may hide implicit assumptions and context-specific knowledge gaps, leading to an overestimation of their usefulness for decision-making. This is especially concerning in safety-critical areas, like critical infrastructure, where risk management decisions are closely tied to national security. Without an explicit epistemic context, risk statements produced by LLMs do not provide enough foundation for risk-based decisions.

Keywords: *Epistemic Uncertainty; Knowledge Strength; LLM-Supported Risk Management; Risk Theory; Decision Support*

INTRODUCTION

The use of large language models (LLMs) in risk management is growing in security-related organizations. LLMs are increasingly promoted in risk management as tools to improve efficiency and consistency, particularly for handling large volumes of unstructured data. Consequently, LLMs are increasingly integrated into existing risk management processes as supportive tools. Previous studies have mainly concentrated on procedural, organizational, and governance issues, such as human oversight, compliance with standards, and responsibility allocation. The performance of LLMs has been well documented through empirical analysis of clearly defined text-based tasks. Nonetheless, it remains unclear how defensible LLM risk statements are in terms of knowledge, especially under high uncertainty and normative context dependence. The main issue is not whether these statements sound plausible, but whether their underlying knowledge, assumptions, and contextual factors are clear enough for critical evaluation. This question serves as the starting point of this study. While there are established concepts for analyzing epistemic uncertainty and the strength of knowledge in risk research (Sahlin et al. 2021; Ylönen and Aven 2023). These have so far rarely been systematically applied to AI-based risk statements and their implicit knowledge assumptions. In the context of critical infrastructure, decisions on risk management are closely linked to national security, public safety, and the continuity of essential services. In such settings, epistemic uncertainty not only impacts analytical quality but also directly influences the legitimacy and accountability of risk-based decisions. The growing use of LLM-supported risk analyses in critical infrastructure, therefore, raises specific epistemic concerns, as linguistically coherent risk statements might conceal unresolved uncertainties with potentially systemic effects.

Risk management is inherently linked to uncertainty. Risk theory distinguishes between two types of uncertainty: aleatory and epistemic. The latter arises from incomplete knowledge, limited data, model assumptions, and contextual dependencies (Sahlin et al. 2021). Although Sahlin et al. (2021) develops this distinction in the context of Bayesian networks for environmental and risk

analysis, the underlying epistemic logic is generally transferable to model-based risk statements, including those generated by large language models. Classical, expert-based risk management approaches make these uncertainties explicitly visible, for example, through documented assumptions, qualitative assessments, or the deliberate labeling of knowledge gaps (Flage and Askeland 2020; Hardaway and Flage 2025). However, the use of LLMs is fundamentally transforming this practice. Model-generated results are often described as linguistically plausible, consistent, and complete, while the underlying knowledge boundaries, assumptions, or uncertainties remain hidden. Against this backdrop, LLM-supported risk analyses may produce well-formed and seemingly complete knowledge claims, while the conditions for assessing their epistemic viability, explicit assumptions, knowledge limits, and context dependence remain unclear. In safety-critical contexts, this lack of clarity is problematic because risk-based decisions require not just consistent statements but also contextualized knowledge whose boundaries can be identified and examined (Aven and Zio 2018).

This article explores epistemic uncertainty in LLM-based risk management. It aims to analyze the risk statements produced by LLMs not just as technical outputs, but as knowledge claims whose epistemic quality and knowledge strength need careful evaluation. Using risk-theoretical and epistemological concepts, we discuss how LLMs may reduce, transform, or potentially obscure existing uncertainties and the impact this has on the quality of risk-based decisions.

RESEARCH METHODOLOGY

This study follows a qualitative-conceptual research design, based on a multi-stage, theory-driven literature analysis. The goal of the methodology is to systematically identify and examine epistemically central contributions to uncertainty and the robustness of knowledge within risk-based decision-making processes. The analysis begins with the risk-theoretical distinction between aleatory and epistemic uncertainty. The latter is defined as uncertainty regarding the quality, completeness, and reliability of the underlying knowledge base. The focus is on works that explicitly address epistemic uncertainty, knowledge assumptions, or the strength of knowledge. Contributions that solely deal with governance, ethics, regulation, explainable AI, or purely technical performance indicators are intentionally excluded.

The primary identification was conducted using freely accessible scientific databases to ensure transparency and reproducibility. The Directory of Open Access Journals (DOAJ) served as the main source, as it includes quality-assured, peer-reviewed journals. Additionally, an AI-based academic discovery tool (Perplexity) was employed to identify potentially relevant publications that are not systematically captured by traditional keyword searches. This tool was used solely for bibliographic research and did not influence the content analysis, interpretation, or evaluation of the sources. To prevent potential misclassifications or hallucinations, all sources identified by the AI-supported discovery tool were manually reviewed and included in the analysis only if they were clearly peer-reviewed and content-valid. As part of the process, the collected sources underwent a three-stage filtering method. First, a strict exclusion check was performed on all identified sources, using predefined criteria based on the sources themselves. These criteria included the absence of explicit epistemic terms, a purely probabilistic approach to uncertainty, an exclusively technical focus, or a lack of references to risk analysis or decision support. The exclusion check was conducted using a standardized scheme to ensure consistent decisions across all sources.

The classification of remaining sources in stage 2 was based on their epistemic contribution. The goal was to distinguish between epistemically central works and those that are merely supporting. During the analysis, each source was placed into one of three categories, with sources in category C being excluded from the analysis:

- a) Central epistemic contributions
- b) Supporting epistemic contributions, or
- c) Peripheral mentions without much conceptual depth.

In the third stage of analysis, only sources classified as central were subjected to a differentiation test. As part of this process, we assessed whether a source presented an independent epistemic

idea or simply repeated existing arguments. Sources whose contribution was deemed conceptually redundant were excluded. The remaining publications were then categorized based on their primary argumentative role. Each source was assigned a specific function, such as defining fundamentals, highlighting conceptual differences, engaging in epistemic critique, examining AI-specific epistemic implications, or analyzing the decision-related aspects of consequences. This functional classification was not used for further filtering but rather for organizing the literature within the paper's argumentative structure.

The methodology of this work is intentionally designed to emphasize theoretical depth and conceptual clarity. The literature is selected in a targeted, theory-driven manner, focusing on epistemic lines of reasoning. Accordingly, the selection emphasizes contributions with an explicit epistemic basis. This focus is essential to adequately analyze epistemic uncertainty in LLM-based risk assessments. Although the analysis is conceptual and does not focus on a specific sector, it is developed with an emphasis on safety-critical environments, such as critical infrastructure. These settings serve as an analytical reference due to their low tolerance for epistemic opacity and their importance for decision-making related to national security.

RESULTS

The results of the conceptual-epistemic analysis show that using large language models in risk management does not primarily reduce uncertainty, but rather changes the form, location, and visibility of epistemic uncertainty. The identified findings can be grouped into the following result areas.

Analysis of the specialist literature indicates that epistemic uncertainty in LLM-based risk analyses manifests in several, sometimes overlapping, forms. Besides uncertainties stemming from incomplete or limited data, there are specific uncertainties related to the robustness of the underlying knowledge base, the implicit model assumptions, and the context dependence of generated risk statements (Sahlin et al. 2021; Ylönen and Aven 2023). These dimensions do not relate to individual parameters but rather to the epistemic status of the entire risk statement. Traditional, expert-based risk analyses often explicitly identify epistemic uncertainty or describe it qualitatively. In contrast, risk statements generated by LLMs usually appear as complete and coherent knowledge assertions. Epistemic uncertainty, therefore, manifests less as a clearly identifiable knowledge gap and more as uncertainty about how much a statement is based on reliable, contextualized knowledge (Sahlin et al. 2021).

Another result of the analysis is the discovery that LLMs systematically alter how knowledge is represented in risk management. Classical risk analyses highlight uncertainty through explicit assumptions, qualitative assessments, or sensitivity considerations. LLMs produce linguistically consistent and plausible texts in which such epistemic markers are mostly absent (Flage and Askeland 2020). The reviewed literature indicates that this type of representation conceals implicit assumptions and knowledge boundaries, making it difficult to examine epistemic assumptions. This can lead to the misconception that a complete knowledge base exists, even though it is actually limited, contingent, or context-dependent (Milkau 2017). Epistemic uncertainty does not vanish during the process. It simply shifts its form and is no longer openly recognized, but stays as an implicit aspect of the generated knowledge representation.

The analysis highlights several systematic epistemic limitations of LLM-generated risk statements. These include, in particular, hallucinations, where linguistically plausible but epistemically unfounded statements are produced (Li et al. 2025), as well as a strong dependence of the results on prompt structure, context selection, and training data (Kirchhof et al. 2025; Yadkori et al. 2024). Furthermore, several studies indicate that LLMs fail to reliably distinguish between verified knowledge, plausible deductions, and speculative additions (Nguyen-Hien et al. 2025; Zeng et al. 2017). Zeng et al. (2017) developed this perspective in the context of the reliability engineering of physical systems. The epistemic heterogeneity discussed in other technical fields can be conceptually applied to LLM-generated knowledge claims. Risk statements can thus be epistemically diverse even if this diversity is not visible in their presentation. Epistemic uncertainty, therefore,

relates not to individual parameters but to the overall epistemic status of the generated statement (Iovane 2025). The uncertainty model developed by Iovane (2025) was originally intended for formal uncertainty analyses in other technical fields, especially in quantum and cyber systems. In this context, it is intentionally used as an analytical analogy to describe the epistemic properties of LLM-generated risk statements. Hardaway and Flage (2025) demonstrate that assumptions in complex, interactive systems are temporarily unstable and continually shift in response to changing contextual conditions. This further complicates the assessment of the epistemic validity of seemingly consistent model outputs.

The results of the previous analysis indicate that the strength of knowledge in risk statements generated by LLMs cannot be determined solely by their linguistic elaboration, internal coherence, or level of detail. In epistemic risk research, the strength of a risk statement is instead defined by factors such as the explicitness of assumptions, the quality and scope of the knowledge base, and the transparent consideration of epistemic uncertainty (Sahlin et al. 2021; Ylönen and Aven 2023). In this context, LLM-based risk analyses create a consistent gap between perceived and actual knowledge. While traditional, expert-based analyses often make uncertainties clear and enable critical assessment, LLMs present risk claims as logical knowledge statements without revealing underlying assumptions, data gaps, or contextual factors (Flage and Askeland 2020; Milkau 2017). This can lead to an overestimation of the epistemic validity of such statements. From an epistemological standpoint, this marks a shift in how epistemic uncertainty is assessed. It moves from explicitly stated assumptions to implicit properties of the model-generated knowledge representation, making it more difficult to evaluate directly (Aven and Zio 2018). The knowledge strength of LLM-generated risk statements is therefore limited and harder to recognize and verify. For evaluating and using these statements in risk management, it is essential to clearly separate linguistic quality from epistemic robustness. The epistemic differences between traditional risk analyses and LLM-generated risk statements identified in the analysis are summarized in Table 1.

Table 1. Epistemic characteristics of classical risk analyses and LLM-generated risk statements

Dimension	Classical risk analysis	LLM-supported risk statements
Type of statement	Explicitly justified risk assessment	Linguistically coherent knowledge claim
Dealing with assumptions	Documented, partially explicit	Implicit, rarely stated
Presentation of epistemic uncertainty	Visible (qualitative/narrative)	Latent, not marked
Context dependency	Partially reflected	Strong, but hidden
Knowledge strength	Can be assessed based on assumptions and data	Difficult to assess
Relevance to decision-making	Transparently limited	Risk of overestimation

DISCUSSION

The analysis shows that using large language models in risk management is not just a technological upgrade of existing methods. Instead, it causes an epistemic shift that changes basic ideas about knowledge, uncertainty, and how risk statements are judged. The following discussion classifies this shift theoretically and explores its implications.

At the core of this epistemic shift is the altered role of epistemic uncertainty. The results indicate that LLMs do not decrease uncertainty but instead change how it appears. While epistemic uncertainty is often explicitly addressed in traditional risk management, such as through documented assumptions or qualitative evaluations, it becomes less prominent in LLM-based analyses (Flage and Askeland 2020; Ylönen and Aven 2023). This shift does not alter the level of epistemic uncertainty but rather its visibility and assessability. Although the epistemic validity of LLM-based risk statements cannot be determined solely from their linguistic coherence, such statements often seem consistent and complete. Consequently, epistemic uncertainty becomes an implicit aspect of the generated knowledge: it is no longer openly expressed through explicit assumptions or appar-

ent knowledge gaps, but is subtly embedded in the output, making it difficult for decision-makers to detect (Hardaway and Flage 2025). Kirchhof et al. (2025) argue that the classic distinction between aleatory and epistemic uncertainty has limited explanatory value for interactive LLM agents. This is because uncertainty here stems not only from randomness or incomplete knowledge but also increasingly from the dynamic interaction among the model, the context of use, and user behavior.

In this context, the importance of knowledge strength as an independent analytical dimension becomes clear. The results support the hypothesis that linguistic quality, internal consistency, or depth of detail of LLM-generated risk statements are not reliable indicators of the strength of the underlying knowledge base (Sahlin et al. 2021). The term “knowledge strength” refers to the justificatory power of a risk statement, taking into account the clarity of its assumptions, the quality and breadth of the underlying knowledge base, model limitations, and contextual factors. Importantly, LLM-based risk analyses usually do not make these preconditions clear; epistemic uncertainty therefore remains relevant for decision-making but tends to be under-indicated systemically. This observation complements existing risk theory work that highlights knowledge strength as a key factor in evaluating risk statements (Aven and Zio 2018).

The identified epistemic shifts directly affect risk-based decision-making processes. Decisions in risk management are not solely based on the availability of consistent information. Instead, the key factor is the ability to recognize the limits of available knowledge and intentionally include uncertainty in decision-making (Aven 2016). There is a risk that LLM-generated risk statements will be seen as more epistemically solid than they actually are because of the epistemic shifts shown in Table 1. In safety-critical situations, this can cause a consistent overestimation of their importance. These results indicate that LLM-generated risk statements without additional epistemic context do not provide a sufficient basis for risk-based decisions (Aven and Zio 2018). At the same time, it should be noted that using large language models can help reduce certain types of uncertainty under specific boundary conditions. In particular, when systematically aggregating large, diverse text corpora and consistently applying predefined analysis and documentation methods, LLMs can structure operational uncertainties and improve the traceability of risk analyses. Risk theory literature indicates that enhanced information processing and more explicit documentation of assumptions can influence how uncertainty is perceived and managed without necessarily altering its epistemic foundations (Aven 2016; Flage and Askeland 2020). However, these effects mainly concern the organizational and procedural aspects of risk analysis, not the epistemic validity of the underlying knowledge base. The current analysis is therefore not aimed at the instrumental usefulness of LLMs, but rather at an epistemic overinterpretation of risk statements that are linguistically plausible but context- and knowledge-limited. This epistemic change poses significant challenges, especially in critical infrastructure settings. In situations characterized by systemic interdependencies and limited reversibility of decisions, epistemic overconfidence in LLM risk assessments could worsen the spread of negative outcomes.

The analysis clearly distinguishes this approach from existing governance, control, and explainability methods. While these methods aim to enhance transparency, traceability, and accountability in AI systems, they do not address the epistemic status of the knowledge claims produced by these systems (Ylönen and Aven 2023). The results indicate that epistemic uncertainty arises from an upstream problem dimension that cannot be resolved solely through explainability or control mechanisms. Questions of control and responsibility assume that the epistemic validity of risk statements can be recognized and assessed in the first place. (Flage and Askeland 2020).

From a conceptual standpoint, this work advances understanding by explicitly treating LLM-generated risk statements as epistemic knowledge claims and connecting their assessment to criteria of knowledge strength. This introduces a theoretical perspective that extends beyond purely technical performance assessments and highlights epistemic uncertainty as a key analytical factor in LLM-supported risk management. At the same time, the work is intentionally structured as a conceptual analysis and does not aim to empirically validate the identified effects. The focus is on epistemic structures and evaluation processes, not on specific decision-making practices or

organizational implementation issues. This limitation defines the scope of the findings and also provides a systematic foundation for future empirical research (Aven 2016).

CONCLUSIONS

The study explored epistemic uncertainty in risk management through a conceptual lens using large language models. It began with the observation that these models are increasingly used to support risk management processes, but the epistemic quality of the risk statements they produce has not been adequately examined.

This study shows that adopting LLMs in risk management requires an epistemic reevaluation of established decision practices. The main challenge is not reducing uncertainty but changing how it is understood: model-generated risk statements often seem clear, complete, and linguistically plausible, while assumptions, knowledge limits, and dependence on context remain implicit (Flage and Askeland 2020; Sahlin et al. 2021). As a result, a systematic gap develops between the perceived quality of LLM-generated risk statements and their actual epistemic viability. Linguistic coherence and formal consistency are therefore not reliable indicators of the underlying knowledge strength, which remains an independent and decision-relevant factor in risk assessment (Aven and Zio 2018; Sahlin et al. 2021). Without explicit epistemic context, such statements do not provide a sufficient basis for risk-based decisions, especially in safety-critical situations, where hidden epistemic uncertainty can lead to a systematic overestimation of decision relevance (Aven 2016; Aven and Zio 2018). This is especially important for risk management in critical infrastructures. Decisions are made within national security contexts in these cases. They also involve high risks, public accountability, and limited tolerance for epistemic ambiguity. In such environments, it is crucial to make epistemic uncertainty and the strength of knowledge explicit to ensure responsible AI-supported risk governance.

This work's contribution lies in clearly classifying LLM-supported risk statements as knowledge claims and connecting this view to the idea of knowledge strength. As a result, the focus shifts from assessing technical performance or governance to examining the epistemic strength of the risk statements themselves (Ylönen and Aven 2023). Regarding future research projects, several starting points emerge, especially in operationalizing epistemic transparency and incorporating concepts of knowledge strength into LLM-supported risk management processes. Empirical studies on how epistemically contextualized risk statements affect decision quality represent a clear next step.

REFERENCES

- AVEN, T., 2016. Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, vol. 253, issue 1, pp. 1–13.
- AVEN, T. and ZIO, E., 2018. *Knowledge in Risk Assessment and Management*. Hoboken, NJ: John Wiley & Sons Inc.
- FLAGE, R. and ASKELAND, T., 2020. Assumptions in quantitative risk assessments: When explicit and when tacit? *Reliability Engineering & System Safety*, vol. 197, 106799.
- HARDAWAY, K. and FLAGE, R., 2025. A framework for evolving assumptions in risk analysis. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, vol. 45, issue 8, pp. 2232–2242.
- IOVANE, G., 2025. An Extended Epistemic Framework Beyond Probability for Quantum Information Processing with Applications in Security, Artificial Intelligence, and Financial Computing. *Entropy*, vol. 27, issue 9.
- KIRCHHOF, M., KASNECI, G. and KASNECI, E., 2025. Position: Uncertainty Quantification Needs Reassessment for Large Language Model Agents. In: *Proceedings of the 42nd International Conference on Machine Learning*, vol. 267, pp. 81665–81677. *Proceedings of Machine Learning Research (PMLR)*. [viewed 31 August 2026]. Available from: <https://proceedings.mlr.press/v267/kirchhof25b.html>
- LI, Z., YI, W. and CHEN, J., 2025. Beyond Accuracy: Rethinking Hallucination and Regulatory Response in Generative AI. *arXiv preprint arXiv:2509.13345*. [viewed 31 August 2026]. Available from: <https://arxiv.org/abs/2509.13345> DOI: <https://doi.org/10.48550/arXiv.2509.13345>
- MILKAU, U., 2017. Risk Culture during the Last 2000 Years – From an Aleatory Society to the Illusion of Risk Control. *International Journal of Financial Studies*, vol. 5, issue 4, p. 31.

- NGUYEN-HIEN, T. D., IVANOVA, D. R., TEH, Y. W. and LEE, W. S., 2025. Extending Epistemic Uncertainty Beyond Parameters Would Assist in Designing Reliable LLMs. *arXiv preprint arXiv:2506.07448*. [viewed 31 August 2026]. Available from: <https://arxiv.org/abs/2506.07448> DOI: <https://doi.org/10.48550/arXiv.2506.07448>
- SAHLIN, U., HELLE, I. and PEREPOLKIN, D., 2021. “This Is What We Don’t Know”: Treating Epistemic Uncertainty in Bayesian Networks for Risk Assessment. *Integrated Environmental Assessment and Management*, vol. 17, issue 1, pp. 221–232.
- YADKORI, Y. A., KUZBORSKI, I., GYÖRGY, A. and SZEPESVÁRI, C., 2024. To Believe or Not to Believe Your LLM: Iterative Prompting for Estimating Epistemic Uncertainty. In: *Proceedings of the 38th International Conference on Neural Information Processing Systems*, pp. 58077–58117. Red Hook, NY: Curran Associates Inc.
- YLÖNEN, M. and AVEN, T., 2023. A framework for understanding risk based on the concepts of ontology and epistemology. *Journal of Risk Research*, vol. 26, issue 6, pp. 581–593.
- ZENG, Z., KANG, R., WEN, M. and ZIO, E., 2017. A Model-Based Reliability Metric Considering Aleatory and Epistemic Uncertainty. *IEEE Access*, vol. 5, pp. 15505–15515.

**ОТ ОЦЕНКА НА РИСКА ДО ТВЪРДЕНИЯ ЗА ЗНАНИЯ:
ЕПИСТЕМНА НЕСИГУРНОСТ В УПРАВЛЕНИЕТО НА РИСКА
ЗА КРИТИЧНИ ИНФРАСТРУКТУРИ, БАЗИРАНО НА ГЕМ (LLM)**

Резюме: Все по-честото използване на големи езикови модели ГЕМ (LLM) в управлението на риска се обосновава с повишена ефективност и по-добра обработка на неструктурирани данни, докато епистемичното качество на получените рискови изявления остава недостатъчно изследвано. Настоящата статия анализира епистемичната несигурност в управлението на риска, базирано на LLM, от концептуална перспектива, опирайки се на теорията на риска и епистемологията. Рисковите изявления, генерирани от LLM, се разглеждат като твърдения за знания и се оценяват по отношение на своята епистемична жизнеспособност и сила на знанието.

Анализът показва, че LLM не намаляват епистемичната несигурност, а променят нейната видимост и възможност за оценка. Плавните и консистентни формулировки могат да прикриват неясни предположения и контекстуални дефицити в знанието, което води до надценяване на тяхната приложимост при вземането на решения. Това е особено критично в области с високи изисквания за безопасност, като критичната инфраструктура. Без изрично епистемично контекстуализиране рисковите изявления, генерирани от LLM, не представляват достатъчна основа за решения, базирани на риска.

Ключови думи: епистемична несигурност; сила на знанието; управление на риска, подкрепено от LLM; теория на риска; подкрепа при вземането на решения

Себастиан Грюнеберг, докторант

ORCID ID: 0009-0002-2339-0729

Университет по библиотекознание и информационни технологии

София, България

E-mail: Sebastian.Grueneberg@outlook.com