# Reducing Hallucinations in LLMs: A Semantic Approach for Cultural Heritage

Fabio Clarizia[1], Massimo De Santo[2], Rosario Gaeta[3], Roberta Gimigliano[4], Rocco Loffredo[5]

***Abstract*—Large Language Models (LLMs) are prone to hallucinations when dealing with specialized domains like cultural heritage, which Retrieval Augmented Generation (RAG) can help mitigate using structured data. This study proposes integrating ontologies into a RAG framework to enhance the semantic accuracy and contextual relevance of LLM outputs. A case study based on artworks at the University of Salerno demonstrates that ontology-based retrieval significantly reduces hallucinations and improves factual alignment. The results suggest that ontology-enhanced LLMs can foster more reliable and meaningful interactions in cultural heritage information systems.**

***Index Terms*—large language models, retrieval augmented generation, ontologies, cultural heritage**

## I. Introduction

The advancement of increasingly powerful Large Language Models (LLMs), such as GPT-4[1] and Claude 4[2], is demonstrating remarkable improvements in both text generation and comprehension, enabling those models to accomplish progressively more complex tasks [1]. However, their performance fundamentally relies on two key factors: the recognition of linguistic patterns through probabilistic inference and the knowledge embedded during the training phase, which define the model's parameters.

This reliance becomes problematic when dealing with queries in highly specialized or localized knowledge domains, such as academic research or region-specific contexts where the models are prone to hallucinations. These are outputs that, while linguistically coherent, lack scientific validity and are often the result of opaque reasoning processes [2].

To improve LLMs for domain-specific applications, current state-of-the-art approaches include full or partial re-training of model parameters. Despite their effectiveness, these methods are computationally expensive, require significant technical expertise, and contribute to substantial $CO_2$ emissions due to the energy required for large-scale training [3].

Alternative strategies involves Prompt Engineering [4], which focuses on writing prompts to guide the model toward more accurate responses. However, one of the most promising and widely adopted approaches in both literature and practice is Retrieval Augmented Generation (RAG) [5]. This technique introduces an auxiliary component, the Retriever, that, upon receiving a user query, retrieves relevant information from an external, often purpose-built, knowledge base. The system then processes an enriched prompt that includes the original query, specific instructions, and the retrieved contextual data and send it as input to the LLM.

Compared to model re-training, RAG offers a more efficient and environmentally sustainable solution [6]. It also enhances transparency, as the supporting information used to generate the response can be traced, an especially valuable feature when working with black-box LLMs, where the internal reasoning is otherwise inaccessible.

Despite the innovation of this method, the effectiveness of RAG method is highly dependent on the quality of the external knowledge base. Unstructured data, i.e. general text, may introduce biases or ambiguities deriving from the authors, which can negatively impact LLM performance. For this reason, it is crucial to construct knowledge bases by using data that are as clear, formal, and domain-specific as possible [7]. By approaching this issue, this research explores the use of structured data, specifically ontologies, which are formal and systematic representations of domain knowledge, as a potential solution for improve LLM specialization.

## II. Methodology

Ontologies can be used as a semantic structured data that can significantly improve the performance of LLMs, particularly in specialized domains such as cultural heritage. An ontology represents knowledge through a structured network of entities (e.g., individuals, locations, artifacts) and their relationships (e.g., createdBy, locatedIn, hasAuthor), mainly encoded as subject–predicate–object triples. Unlike unstructured text, ontologies clarifies and organizes conceptual relationships, making them explicitly navigable. Integrating such structured knowledge into RAG frameworks can lead to outputs that are more factually accurate, contextually appropriate, and semantically coherent within the target domain [7], [8].

This research is co-financed by Netcom Engineering S.p.A.

[1]Fabio Clarizia is with the Department of Humanities, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA, Italy, fclarizia@unisa.it

[2]Massimo De Santo is with the Department of Industrial Engineering, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA, Italy, desanto@unisa.it

[3]Rosario Gaeta is with the Department of Industrial Engineering, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA, Italy, rgaeta@unisa.it

[4]Roberta Gimigliano is with the "P.S. Mancini" State Scientific High School, Via De Concilii, 1, 83100 Avellino AV, Italy, roberta.gimigliano@gmail.com

[5]Rocco Loffredo is with the Department of Industrial Engineering, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA, Italy, rloffredo@unisa.it

[1]https://openai.com/chatgpt/overview/

[2]https://claude.ai/

In the cultural heritage domain, where data are complex and accuracy is fundamental, recent research has begun to explore the use of ontologies as external knowledge bases for LLMs to improve contextual relevance and accuracy of generated content [9]. This starting point open doors to further and encouraging roads for enhancing AI-driven interpretation of heritage materials and reducing the risk of misrepresentation or misinformation. Moreover, this approach is completely opposite from the prevailing trend in the research field, which focuses on using LLMs to generate, refine, or update ontologies themselves [10].

The proposed framework relies in the semantic embedding of both the domain ontology and the user's natural language input. Once embedded into a shared vector space, similarity-based retrieval techniques, such as cosine similarity, are employed to identify and rank the most relevant knowledge elements in response to the user's query by making semantic comparison.

One of keys advantage of this approach is the flexibility regarding the source of the domain ontology. Ontologies can be derived from manual realization by domain-experts and from total or semi-automatic methods from structured data [11] or generated from unstructured content [12], leveraging existing ontology learning tools that support automatic or semi-automatic extraction from heterogeneous data sources.

This framework enables a retrieval process that is less hallucinated and more context-aware than traditional keyword-searching-based methods by capturing semantic relationships and hidden connections between concepts. The synergy between structured semantic knowledge and the generative capabilities of LLMs lays the basis for a more intelligent, transparent, and trustworthy information retrieval paradigm.

## III. Case of Study

To evaluate the proposed framework, a case study was conducted using an ontology based on the artworks within the University of Salerno. Following the embedding of the ontology and the deployment of the RAG system, a set of domain-specific questions was formulated to test the system's performance in terms of faithfulness and relevancy of the answers. These same questions were also submitted to general-purpose and agnostic LLMs, which were provided with relevant contextual information, to benchmark the ability to generate accurate and contextually appropriate responses.

The primary aim of this experimental setup was to make a comparison between the performance of information retrieval from two different knowledge bases composed of structured (ontological) and unstructured data sources.

First results show that structured data retrieval significantly improves the accuracy of the generated responses. In particular, answers derived from the structured ontology demonstrated a higher degree of alignment with expert-validated ground truth, exhibited fewer hallucinations rate, and maintained closer relevance to the questions.

## IV. Conclusion

Ontology-enhanced LLMs show a promising approach to enhance the reliability of cultural heritage information systems. Unlike traditional keyword-search-based systems or static database queries, an LLM augmented with a domain-specific knowledge graph enables more dynamic, semantically enriched, and context-aware interactions with heritage data.

This methodology capitalizes on the generative strengths of LLMs while anchoring their outputs in structured, validated knowledge sources, therefore improving both factual precision and user trust. Preliminary experiments confirm that even the incorporation of relatively simple ontologies can substantially reduce hallucinations and factual inaccuracies in generated responses.

This work offers future research and practical implementations. With advancements in semantic embedding techniques and robust evaluation methodologies, this framework holds the potential to evolve digital cultural heritage systems, offering more accurate, meaningful, and engaging experiences for both experts and the broader public.

## References


[1] X. Tu, Z. He, Y. Huang, Z.-H. Zhang, M. Yang, and J. Zhao, "An overview of large ai models and their applications," *Visual Intelligence*, vol. 2, no. 1, 2024, cited by: 5.

[2] Z. Xu, S. Jain, and M. Kankanhalli, "Hallucination is inevitable: An innate limitation of large language models," *arXiv preprint arXiv:2401.11817*, 2024.

[3] A. Faiz, S. Kaneda, R. Wang, R. Osi, P. Sharma, F. Chen, and L. Jiang, "Llmcarbon: Modeling the end-to-end carbon footprint of large language models," 2024. [Online]. Available: https://arxiv.org/abs/2309.14393

[4] G. Marvin, N. Hellen, D. Jjingo, and J. Nakatumba-Nabende, "Prompt engineering in large language models," in *International conference on data intelligence and cognitive informatics*. Springer, 2023, pp. 387–402.

[5] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," *Advances in neural information processing systems*, vol. 33, pp. 9459–9474, 2020.

[6] J. S. Erickson, H. Santos, V. Pinheiro, J. P. McCusker, and D. L. McGuinness, "Llm experimentation through knowledge graphs: Towards improved management, repeatability, and verification," *Journal of Web Semantics*, vol. 85, p. 100853, 2025.

[7] R. S. Dehal, M. Sharma, and E. Rajabi, "Knowledge graphs and their reciprocal relationship with large language models," *Machine Learning and Knowledge Extraction*, vol. 7, no. 2, p. 38, 2025.

[8] E. Lavrinovics, R. Biswas, J. Bjerva, and K. Hose, "Knowledge graphs, large language models, and hallucinations: An nlp perspective," 2024. [Online]. Available: https://arxiv.org/abs/2411.14258

[9] R. Loffredo and M. De Santo, "Using ontologies for llm applications in cultural heritage," in *Proceedings of the 3rd Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2024) co-located with the 23rd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2024)*, 2024.

[10] V. K. Kommineni, B. König-Ries, and S. Samuel, "From human experts to machines: An llm supported approach to ontology and knowledge graph construction," *arXiv preprint arXiv:2403.08345*, 2024.

[11] F. Clarizia, M. De Santo, R. Gaeta, and R. Mosca, "Method for ontology learning from an rdb: Application to the domain of cultural heritage," in *International Conference on Image Analysis and Processing*. Springer, 2023, pp. 409–421.

[12] S. Elnagar, V. Yoon, and M. A. Thomas, "An automatic ontology generation framework with an organizational perspective," 2022. [Online]. Available: https://arxiv.org/abs/2201.05910