Modeling Historical Color Knowledge Leveraging LLMs, Ontologies, and Knowledge Graphs

Mohamed Wadhah Mabrouk[a,*]

[a]Laboratoire d'Informatique de Grenoble - LIG

Abstract. Understanding the multifaceted knowledge surrounding colors in the 15th and 16th centuries requires addressing both their material and symbolic dimensions. This dual perspective is vividly preserved in historical manuscripts, from alchemical treatises to heraldic documents that describe color-related know-how intertwining chemistry, symbolism, culture, and aesthetics. However, because this knowledge is expressed in natural language that is often old-world and context-dependent, it remains difficult to process, structure, or query computationally.

This doctoral research tackles the challenge of representing this complex knowledge leveraging artificial intelligence techniques, in particular Large Language Models (LLMs) guided by structured templates and semantic ontologies. Our aim is to enable the systematic extraction, representation, and reasoning over historical knowledge about colors both as material substances (e.g., pigments, compounds) and as semantic signifiers (e.g., heraldic meanings, alchemical powers).

1 Introduction

Understanding how pre-modern craftsmen, alchemists, and artists worked with color means reconstructing their recipes not just as static lists of ingredients but as *dynamic processes*: sequences of operations (grind, infuse, boil, precipitate) that unfold under particular agents such as the four elements (water, fire, earth, and air) and in which each substance can change status from a raw pigment source to an intermediate preparation or the chromatic goal itself. Scholars therefore need a framework that can:

- Distinguish raw materials from transformed ones.
- Pinpoint the alchemical actions [12, 20] applied to them.
- Identify the evolving *function* of each colour term (ingredient, intermediate, final hue, symbolic marker, chemical species).
- Capture the "shape" or *form* of a recipe or treatise—that is, its rhetorical scaffolding, from lists of requisites and sequential imperatives to embedded moral or symbolic digressions.

To operationalize and evaluate such a framework, we based our study on the *Patrimalp*[1]. corpus: a curated collection of roughly 30 historical works that bring together color-related recipes and treatises in medicine, gemstone fabrication (both natural and artificial), cosmetics, heraldry, textiles, and more. Several texts belong to alchemical [4] or magical traditions, where interpretation is further complicated by esoteric or allegorical language, making the corpus an ideal proving ground for methods capable of disentangling material, procedural, and symbolic meanings.

This constraint drives our research, which explores effective methodologies that combine Large Language Models (LLMs) with structured knowledge-representation techniques. Concretely, we design and build an original interdisciplinary knowledge graph from the ground up, then automatically populate it with information extracted by LLMs guided by a procedural modeling language. The resulting graph bridges colorants as physical and chemical entities, their symbolic meanings contextualized by domain and period, their discursive functions in historical recipes, and their diverse historical usages.

2 Related Work

However, there is still little research that directly connects LLMs with the extraction of knowledge related to colors; some studies nonetheless underscore the potential of Large Language Models (LLMs) to enhance cross-linguistic analysis of color semantics, as demonstrated by Song and Wang [17] in their BERT-based study of the term "red" in Chinese and English. A related effort is presented in a recent prototype by Japanese researchers [10], where LLMs are employed to extract semantic knowledge—including color-related terms—from ancient Japanese texts, highlighting the relevance of LLM-driven methods in cultural heritage domains.

Other studies have further advanced LLM integration into ontology and knowledge graph construction, leveraging competency questions, ontology reuse, and RAG for semi-automatic KG population [13]. These works demonstrate LLM effectiveness in structured knowledge extraction [2, 3], ontology enrichment and population [6, 15], and prompt-based entity and relation extraction [**?**]. However, challenges like prompt sensitivity, hallucinations, and validation persist. Emerging hybrid frameworks also suggest mutual support between knowledge graphs and LLM reasoning [11, 1]. Despite these advances, reproducibility, interoperability, and evaluation remain significant challenges, especially in interdisciplinary contexts such as ours, involving symbolic, material, and procedural knowledge dimensions. Our work addresses these gaps by integrating structured extraction guided by LLMs with layered ontologies and expert validation to ensure semantic depth and historical accuracy.

3 Research Questions

This research is guided by the following questions:

- **RQ1.** How can we extract structured, semantically valid knowledge from ancient texts using LLMs without introducing hallucinations or misinterpretations?

* Corresponding Author. Email: mohamed.mabrouk1@univ-grenoble-alpes.fr

[1] https://patrimalp.univ-grenoble-alpes.fr/

- **RQ2.** What knowledge representation models are suitable to represent both the procedural nature of manuscripts and the symbolic, material, and chemical dimensions of colors?
- **RQ3.** How can an interdisciplinary ontology unify historical, symbolic, and chemical dimensions while remaining interoperable with existing ontologies?
- **RQ4.** How can we enable temporal and semantic reasoning across centuries?
- **RQ5.** What are the limitations of current LLMs in modeling procedural and symbolic knowledge, and how can we evaluate their outputs?

4 Statement of Thesis

This research adopts a methodology that combines neural and symbolic AI to address the challenges of representing historical color-related knowledge across procedural, symbolic, and chemical dimensions.

To extract procedural knowledge from historical color texts, we adopt a structured pipeline centered on a concise *task-based knowledge model*, **TaskLang** (inspired from AromTask [16]). Prior evaluations have shown that while LLMs can generate entity-relation triples [14], they often lack ontological rigor and struggle with faithful process extraction: they may skip steps, merge branches, or invent activities not present in the text [19, 8]. We therefore propose anchoring the LLM's reasoning to this task model, which frames each expected activity as an explicit schema-guided task, in order to curb hallucination and yield semantically sound, machine-readable process models.

Historical texts are thus first interpreted by LLMs using customized prompts [9], ensuring syntactic and semantic consistency while minimizing the risk of hallucinations (**RQ1**, **RQ5**). The extracted content is then transformed into structured representations following the TaskLang structure, capturing goals, inputs, actions, tools, and dependencies. These outputs are further converted into RDF schemas and stored in a Neo4j database, enabling visual inspection, semantic querying, and integration with other knowledge layers.

To unify the extracted procedural knowledge with broader conceptual interpretations, we are developing an interdisciplinary ontology called *Coloronto* following established ontology engineering methodologies [5]. It supports multiple layers of color-related knowledge structured across distinct yet interrelated domains, including alchemy, chemistry, heraldry, religious symbolism, medicine, artistic practices, orfèvrerie, and cosmetic recipes. Coloronto models the polyvalent nature of colors by capturing both their material composition and their interpretive roles in historical contexts. Notably, color-related entities extracted via TaskLang (e.g., ingredients, materials, target hues) are systematically linked to corresponding concepts in Coloronto, enabling semantic enrichment and supporting both symbolic and chemical interpretations of procedural knowledge (**RQ2**, **RQ3**).

The ontology currently incorporates:

- **Colorants** as physical and chemical entities,
- **Symbolic meanings** contextualized by domain, culture, and historical period,
- **Discursive functions** within textual forms such as recipes or treatises,
- **Historical usages** including regions, textual genres, and transmission paths,

Semantic interoperability is ensured via ontological alignment with existing reference domain ontologies [7] (using constructs such as `owl:sameAs`, `skos:exactMatch`, and contextual properties like `usedInPeriod`, `hasSymbolism`, and `hasAlternative`). Initial alignments can, for example, link traditional pigments (e.g., *cinabre*) to formal chemical identifiers such as `chebi:51102`[2].

This work does not stop at the construction of a knowledge base. In future stages, we will exploit the resulting structured data within a *MatchMaker*—a semantic engine designed to identify, align, and compare color-related concepts across time periods. The MatchMaker will allow us to trace the diachronic evolution of symbolic, material, and functional dimensions of colors. For instance, the symbolic evolution of the heraldic color *gueules* is represented through temporally contextualized annotations (e.g., red as a mark of nobility in 15th-century heraldry, or martyrdom in 16th-century religious texts), enabling semantic queries on meaning shifts across centuries (**RQ4**) [18].

5 Contributions So Far

Preliminary results have been obtained along both axes of the proposed research framework.

Procedural modeling and extraction:
We have implemented a semi-automatic pipeline that guides LLMs (GPT-4o) using TaskLang templates to extract structured procedural knowledge from historical color recipes. To date, over 20 recipes from 15th and 16th-century treatises have been processed, resulting in formalized task representations that capture goals, actions, instruments, and inputs. These representations are serialized in RDF and integrated into a Neo4j graph database for visual exploration and semantic querying.

Ontology design and alignment:
The core concepts of the Coloronto ontology, which cover the symbolic, material, and functional dimensions of colors as found in our corpus, have been mostly defined at this stage.

Exploratory modules and evaluation:
A prototype of a "MatchMaker" module has been initiated to semantically compare the evolution of color meanings across time, using a RAG (Retrieval-Augmented Generation) architecture over the extracted data. This module is being explored as a means to assess how LLMs can support historical reasoning across symbolic and material contexts. In parallel, work has begun on designing a gold standard to evaluate the accuracy and consistency of the LLM-based extraction pipeline.

6 Conclusion

In conclusion, this research contributes to demonstrating the viability of integrating Large Language Models (LLMs) with structured templates and interdisciplinary ontologies for modeling historical knowledge related to colors. By addressing symbolic, procedural, and chemical dimensions within a unified semantic framework, this hybrid approach helps bridge gaps in interdisciplinary knowledge representation and reasoning. Beyond the specific case of color knowledge, the proposed methodology combining LLM-guided extraction with ontology-based structuring may also **be adapted** to other cultural heritage and scientific domains where historical texts must be disentangled, structured, and semantically enriched, such as medicine, material sciences, or the history of technology.

[2] https://www.ebi.ac.uk/chebi/

References

[1] S. Carta, A. Giuliani, L. Piano, A. S. Podda, L. Pompianu, and S. G. Tiddia. Iterative zero-shot llm prompting for knowledge graph construction. *arXiv preprint arXiv:2307.01128*, 2023.

[2] J. Dagdelen, A. Dunn, S. Lee, N. Walker, A. S. Rosen, G. Ceder, K. A. Persson, and A. Jain. Structured information extraction from scientific text with large language models. *Nature Communications*, 15(1418), 2024. doi: 10.1038/s41467-024-45563-x. URL https://doi.org/10.1038/s41467-024-45563-x.

[3] A. Dunn, J. Dagdelen, N. Walker, S. Lee, A. S. Rosen, G. Ceder, K. Persson, and A. Jain. Structured information extraction from complex scientific text with fine-tuned large language models. 2022. URL https://arxiv.org/abs/2212.05238.

[4] L. Fabbrizzi. Communicating about matter with symbols: Evolving from alchemy to chemistry. *Journal of Chemical Education*, 85(11): 1501–1511, Nov. 2008. doi: 10.1021/ed085p1501.

[5] M. Fernández-López, A. Gómez-Pérez, and N. Juristo. Methontology: From ontological art towards ontological engineering. *AAAI Technical Report SS-97-06*, pages 33–40, 1997. Available at https://www.researchgate.net/publication/50236211.

[6] M. Funk, S. Hosemann, J. C. Jung, and C. Lutz. Towards ontology construction with language models, 2023. URL https://arxiv.org/abs/2309.09898.

[7] A. Gangemi. Ontology reuse: the real test of ontological design. *https://ebooks.iospress.nl/DOI/10.3233/FAIA220292*, 2022. URL https://arxiv.org/abs/2205.02892.

[8] S. Ghosh, N. R. Brodnik, C. Frey, C. Holgate, T. M. Pollock, S. Daly, and S. Carton. Toward reliable ad-hoc scientific information extraction: A case study on two materials datasets. *https://aclanthology.org/2024.findings-acl.897/*, 2024.

[9] H. Gupta, R. Xie, and H. Ji. Knowledge-augmented language model prompting: A survey. *arXiv preprint arXiv:2311.02403*, 2023.

[10] T. Hasegawa, Y. Kubota, and K. Mineshima. Generating semantic knowledge from classical japanese texts using large language models. In *Proceedings of the 38th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI 2024)*, pages 2Q5–IS–1–05, 2024. Presented at the JSAI 2024 Interdisciplinary Session.

[11] H. Ji, T. Yu, R. Xie, M. Yasunaga, R. Rony, S. Choudhury, Y. Gao, J. Finnie-Ansley, T. Gao, X. Ren, et al. Unifying large language models and knowledge graphs: A roadmap. *arXiv preprint arXiv:2306.08302*, 2023.

[12] B. Joly. *La rationalité de l'alchimie au XVIIe siècle*. Mathesis. Librairie Philosophique Vrin, Paris, 1992. ISBN 978-2-7116-1055-6.

[13] V. K. Kommineni, B. König-Ries, and S. Samuel. From human experts to machines: An llm supported approach to ontology and knowledge graph construction. 2024. URL https://arxiv.org/abs/2403.08345. arXiv:2403.08345.

[14] C. Meyer, M. Cetto, M. Scholten, A. Jentzsch, and A. Harth. From prompts to knowledge: How chatgpt performs on the knowledge graph construction pipeline. *arXiv preprint arXiv:2307.06917*, 2023.

[15] S. S. Norouzi, A. Barua, A. Christou, N. Gautam, A. Eells, P. Hitzler, and C. Shimizu. Ontology population using llms. *arXiv preprint arXiv:2411.01612*, 2024. URL https://arxiv.org/abs/2411.01612.

[16] M. Page, J. Gensel, C. Capponi, C. Bruley, P. Genoud, and D. Ziébelin. Représentation de connaissances au moyen de classes et d'associations: le système AROM. In *Actes des Journées Langages et Modèles à Objets (LMO 2000)*, pages 91–106, Mont Saint-Hilaire, Québec, Canada, jan 2000. URL https://www.researchgate.net/publication/221656380. Journées LMO, 25–28 janvier 2000.

[17] Y. Song and D. Wang. Bert-assisted behavioral profiling of polysemy: contrastive analysis of HONG in chinese and RED in english. *Corpus Linguistics and Linguistic Theory*. doi: https://doi.org/10.1515/cllt-2024-0029.

[18] S. Sugiarto, L. Zerkowski, O. Glickman, D. Luther, K. Weisel, and A. Mehler. Modeling concept drift for historical research in the digital humanities. In *Proceedings of the 4th Workshop on Ontologies for Digital Humanities and their Social Analysis (WODHSA)*, 2024. URL https://ceur-ws.org/Vol-2518/paper-WODHSA6.pdf.

[19] W. Van Woensel and S. Motie. Nlp4pbm: A systematic review on process extraction using natural language processing with rule-based, machine and deep learning methods. *Enterprise Information Systems*, 2024. URL https://arxiv.org/abs/2409.13738. Preprint submitted on 24 September 2024.

[20] J. von Lennep. *Alchimie: Contribution à l'histoire de l'art alchimique*. Dervy, Bruxelles, 1990. Première édition 1984.