machine learning & knowledge extraction

MDPI

Article

Stitching History into Semantics: LLM-Supported Knowledge Graph Engineering for 19th-Century Greek Bookbinding

Dimitrios Doumanas [1,*], Efthalia Ntalouka [2], Costas Vassilakis [3], Manolis Wallace [4] and Konstantinos Kotis [1,*]

[1] Intelligent Systems Lab, Department of Cultural Technology and Communication, University of the Aegean, 81100 Mytilene, Greece
[2] Department of Conservation of Antiquities and Works of Art, University of West Attica, Ag. Spyridonos Str., Egaleo, 12243 Athens, Greece; entalouka@uniwa.gr
[3] Software and Database Systems Laboratory (SoDa.Lab), Department of Informatics and Telecommunications, University of the Peloponnese, 22100 Tripolis, Greece; costas@uop.gr
[4] Knowledge and Uncertainty Research Laboratory, Department of Informatics and Telecommunications, University of the Peloponnese, 22100 Tripolis, Greece; wallace@uop.gr
* Correspondence: ctd25002@ct.aegean.gr (D.D.); kotis@aegean.gr (K.K.)

Abstract

Preserving cultural heritage can be efficiently supported by structured and semantic representation of historical artifacts. Bookbinding, a critical aspect of book history, provides valuable insights into past craftsmanship, material use, and conservation practices. However, existing bibliographic records often lack the depth needed to analyze bookbinding techniques, provenance, and preservation status. This paper presents a proof-of-concept system that explores how Large Language Models (LLMs) can support knowledge graph engineering within the context of 19th-century Greek bookbinding (1830–1900), and as a result, generate a domain-specific ontology and a knowledge graph. Our ontology encapsulates materials, binding techniques, artistic styles, and conservation history, integrating metadata standards like MARC and Dublin Core to ensure interoperability with existing library and archival systems. To validate its effectiveness, we construct a Neo4j knowledge graph, based on the generated ontology and utilize Cypher Queries—including LLM-generated queries—to extract insights about bookbinding practices and trends. This study also explores how semantic reasoning over the knowledge graph can identify historical binding patterns, assess book conservation needs, and infer relationships between bookbinding workshops. Unlike previous bibliographic ontologies, our approach provides a comprehensive, semantically rich representation of bookbinding history, methods and techniques, supporting scholars, conservators, and cultural heritage institutions. By demonstrating how LLMs can assist in ontology/KG creation and query generation, we introduce and evaluate a semi-automated pipeline as a methodological demonstration for studying historical bookbinding, contributing to digital humanities, book conservation, and cultural informatics. Finally, the proposed approach can be used in other domains, thus, being generally applicable in knowledge engineering.

Keywords: ontology engineering; knowledge graphs; cultural heritage; bookbinding history; large language models

check for updates

Academic Editors: Mourad Abbas, Najim Dehak and Gérard Chollet

Received: 12 May 2025
Revised: 12 June 2025
Accepted: 23 June 2025
Published: 24 June 2025

Citation: Doumanas, D.; Ntalouka, E.; Vassilakis, C.; Wallace, M.; Kotis, K. Stitching History into Semantics: LLM-Supported Knowledge Graph Engineering for 19th-Century Greek Bookbinding. *Mach. Learn. Knowl. Extr.* **2025**, *7*, 59. https://doi.org/10.3390/make7030059

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cultural heritage preservation is a field that seeks to document, safeguard, and transmit historical artifacts, traditions, and knowledge across generations [1]. Among these

artifacts, books hold a unique place as both intellectual repositories and physical objects of craftsmanship [2]. Beyond their textual content, books are material witnesses of historical periods, reflecting the technological, artistic, and socio-economic conditions of their time. Bookbinding, in particular, is an essential yet often overlooked aspect of book history. It provides insight into the materials, decorative traditions, and structural techniques employed by bookbinders, revealing historical patterns of production, trade, and artistic influence [3]. The study of bookbinding is crucial not only for bibliographic and art historical research, but also for conservation and restoration efforts, ensuring that historical books remain accessible and well-preserved for future generations [4,5]. Despite its significance, bookbinding remains significantly underrepresented in traditional library and archival metadata [6]. Systems such as MARC and Dublin Core emphasize bibliographic and textual data—author, title, and publication date—but provide limited capacity for recording detailed physical features such as sewing structures, spine types, cover materials, or decorative styles. As a result, critical dimensions of material history are left undocumented or inconsistently recorded, inhibiting comparative analysis, provenance research, and conservation planning.

In recent years, ontologies have emerged as a powerful tool for structuring and preserving cultural heritage knowledge [1]. Ontologies provide a formalized representation of concepts and their relationships within a domain, enabling more sophisticated data integration, retrieval, and reasoning [7]. In the context of cultural heritage, ontologies facilitate the standardization and interoperability of data across institutions, allowing libraries, museums, and archives to share and analyze information in a semantically rich manner [8]. Unlike traditional metadata models that rely on predefined fields with limited expressiveness, ontologies support complex relationships, enabling scholars and conservators to trace historical patterns, identify stylistic influences, and assess the cultural significance of artifacts with greater precision [9]. The application of ontologies to bookbinding is particularly promising, as it allows for a more structured documentation of materials, binding techniques, artistic styles, and conservation histories, thus bridging the gap between bibliographic cataloging and material studies. However, while ontologies provide the conceptual backbone for describing cultural heritage knowledge, knowledge graphs take this a step further by operationalizing these formal structures into dynamic, interconnected networks of real-world instances [10]. Knowledge graphs instantiate ontological classes and relationships with actual data, allowing for advanced querying, semantic reasoning, and contextual enrichment. In this way, knowledge graphs extend the expressive power of ontologies, enabling not only the preservation of complex cultural information but also its active exploration and intelligent discovery across domains [11].

The integration of Large Language Models (LLMs) to knowledge graph (KG) engineering introduces new possibilities for automating and enhancing the documentation of cultural heritage. LLMs have demonstrated remarkable capabilities in natural language processing, enabling the extraction, structuring, and reasoning over domain-specific knowledge at an unprecedented scale. In the context of ontology development, LLMs can assist in generating taxonomies, identifying relationships between concepts, and formulating complex queries for knowledge retrieval [12,13]. For cultural heritage applications, this means that vast amounts of textual data—ranging from historical records to conservation reports—can be systematically processed and incorporated into structured knowledge systems. By leveraging LLMs in the construction of bookbinding ontologies and knowledge graphs, it becomes possible to automate parts of the ontology engineering process, enhance semantic reasoning, and provide users with intelligent querying capabilities that facilitate research and conservation efforts [13,14]. This approach moves beyond static metadata structures, offering a dynamic, machine-readable representation of bookbinding knowledge that can evolve over time with new data inputs. This case study addresses

that specific gap by demonstrating how a structured, semantically rich knowledge graph can encode these overlooked physical and conservation-related attributes using LLM-assisted modeling.

This study aims to develop a knowledge graph-driven framework for documenting and analyzing 19th-century Greek bookbinding (1830–1900), addressing the limitations of existing bibliographic records. By constructing a domain-specific ontology that encapsulates bookbinding materials, techniques, artistic styles, and conservation histories, we provide a structured and semantically rich representation of historical bookbinding practices. Unlike previous efforts that focus primarily on textual metadata, our approach integrates detailed material and structural descriptions, enabling more comprehensive analysis and cross-referencing between bookbinding elements. This ontology is complemented by a knowledge graph implemented in Neo4j, which allows for advanced querying and pattern recognition using Cypher, including queries generated through LLMs. By doing so, we introduce a scalable and automated methodology for studying bookbinding history, facilitating a deeper understanding of bookbinding trends, conservation needs, and historical workshop relationships.

A key innovation of our research lies in its ability to provide a more nuanced and interconnected understanding of bookbinding history. Existing bibliographic ontologies tend to be limited in scope, often focusing on textual and publication metadata while neglecting the physical aspects of books. Our ontology enriches this landscape by incorporating detailed representations of binding materials, techniques, artistic elements, and restoration histories, which are later translated to a knowledge graph. Additionally, the integration of LLM-assisted query generation enhances usability, allowing researchers, conservators, and librarians to interact with the data more effectively. By supporting semantic reasoning and intelligent search capabilities, our approach enables users to uncover historical patterns, track conservation histories, and identify stylistic connections between different bookbinding traditions, thereby advancing the study of bookbinding as both a material and cultural phenomenon.

The knowledge graph provides significant value to book maintainers by offering a structured framework for cataloging bookbinding features and maintenance records, aiding in preservation and restoration efforts. Through its detailed documentation of bookbinding elements, it allows conservators to assess material integrity, track wear, and damage, and make informed restoration decisions. The knowledge graph also enhances the ability of libraries and archives to maintain detailed records, improving the discoverability and accessibility of historical bookbinding information. By standardizing data entry and supporting interoperability with existing metadata frameworks such as MARC and Dublin Core, it ensures that bookbinding knowledge can be seamlessly integrated into broader library and museum systems, facilitating collaboration between institutions and researchers.

Beyond book maintainers, this study contributes to the broader digital heritage community by proposing an extension to existing cultural informatics frameworks. Our knowledge graph-driven approach aligns with contemporary efforts to move beyond static cataloging towards a more dynamic representation of cultural heritage artifacts. The structured integration of bookbinding knowledge into digital heritage repositories represents a shift toward a more holistic approach to cultural documentation—one that not only records the physical attributes of artifacts, but also tracks their historical transformations, conservation histories, and cultural significance. By extending bookbinding data to platforms such as Wikidata, this knowledge graph has the potential to enhance digital heritage initiatives, providing researchers, conservators, and the public with a richer, more interconnected understanding of historical books as both textual and material objects.

In summary, this paper presents a novel methodology for documenting and analyzing historical bookbinding through the integration of ontology engineering, knowledge graph construction, and LLM-assisted reasoning. By addressing gaps in bibliographic metadata and introducing a more detailed, semantically rich representation of bookbinding history, our study provides a scalable and automated solution for book conservation and research. This work contributes to the fields of digital humanities, book conservation, and cultural informatics, ultimately enriching our understanding of the craftsmanship, historical evolution, and cultural significance of 19th-century Greek bookbinding.

This study does not aim to provide a comprehensive solution to cultural heritage metadata challenges. Instead, it presents a focused case study—a proof of concept—that explores how current LLM capabilities can assist in building semantically rich knowledge graphs from structured heritage data. Our goal is to demonstrate the potential, limitations, and methodological considerations of applying state-of-the-art LLMs in this context, thereby contributing to ongoing discussions around automation and metadata augmentation in digital heritage.

The intended audience for this study includes digital humanities researchers, librarians, conservators, and cultural heritage professionals who may be unfamiliar with advanced semantic technologies but are exploring the potential of LLM-assisted systems. Rather than assuming deep expertise in ontology engineering or graph databases, we present a practical, guided demonstration of how these technologies can be applied in a real-world domain. In doing so, we aim to support both methodological awareness and hands-on experimentation among professionals working with structured heritage metadata.

The structure of the paper is as follows. Section 2 presents related work. Section 3 describes the proposed approach followed for this line of research. Section 4 presents the experiments and their results, and Section 5 discusses them critically. Finally, Section 6 summarizes the key findings of this study and provides plans for future research.

2. Related Work

Moraitou et al. [15] propose an ontology-based framework to support decision-making in conservation and restoration (CnR) interventions for cultural heritage, addressing the complexities of selecting appropriate conservation strategies. Their research introduces the DS-CnRI framework, which formalizes expert knowledge through a structured ontology, enabling conservators to systematically assess intervention options by considering intrinsic and extrinsic requirements. By integrating Semantic Web technologies and reasoning mechanisms, the framework facilitates interoperability and enhances documentation practices. Their evaluation, conducted in collaboration with conservators, demonstrates that the system effectively aids in selecting suitable interventions, identifying limitations of different options, and improving the overall decision-making process. However, the study acknowledges certain limitations, including the need for broader domain coverage beyond the tested conservation cases, potential challenges in integrating diverse conservation datasets, and the reliance on expert input for ontology refinement. These constraints highlight the necessity for continuous expansion and refinement of ontology-driven decision-support tools in cultural heritage conservation.

Koutsiana et al. [16] explore the use of LLMs in knowledge engineering (KE), particularly their impact on KG construction and ontology development. Their study investigates how knowledge engineers integrate LLMs into KE tasks, examining the challenges, required skills, and ethical considerations involved. Through a multi-method study, including an ethnographic hackathon and semi-structured interviews, they find that LLMs can enhance efficiency in KE by automating tasks such as information extraction, ontology alignment,

and knowledge graph enrichment. However, their findings also highlight critical challenges, including inconsistencies in LLM-generated outputs, the difficulty of evaluating the reliability of automated KE processes, and concerns over responsible AI use. The study concludes that while LLMs hold potential as AI copilots in KE, their adoption requires careful oversight, ethical guidelines, and improved evaluation mechanisms to ensure the trustworthiness and accuracy of generated knowledge.

Linxen et al. [17] introduce an ontology-driven knowledge base aimed at enhancing knowledge organization and retrieval at the Folkwang University of the Arts library, addressing the challenges posed by managing diverse collections of digital and physical media. Their research develops an ontology that integrates structured, semi-structured, and unstructured data into a unified knowledge base, enabling more effective semantic search capabilities through SPARQL queries and LLM-generated prompts. The evaluation demonstrates that the ontology enhances knowledge retrieval efficiency, improves accessibility, and supports semantic integration across different data formats. However, the study also identifies limitations, including performance bottlenecks in processing complex queries, occasional inaccuracies in search results, and the need for further optimization to enhance query response times and result precision. The research concludes that while ontology-driven approaches significantly improve academic library systems, ongoing refinements are necessary to address scalability and retrieval accuracy issues.

Maree [18] introduces a neuro-symbolic approach to relational exploration in cultural heritage knowledge graphs, leveraging LLMs for explanation generation and a mathematical framework to quantify the "interestingness" of relationships. The study employs the Wikidata Cultural Heritage Linked Open Data (WCH-LOD) dataset to validate the framework, demonstrating improvements in precision (0.70), recall (0.68), and F1-score (0.69) compared to graph-based (0.26 F1-score) and knowledge-based baselines (0.43 F1-score). The research highlights that LLM-powered explanations achieve higher BLEU (0.52), ROUGE-L (0.58), and METEOR (0.63) scores, underscoring their effectiveness in enhancing relational search. The findings reveal a strong correlation (0.65) between the interestingness measure and explanation quality, validating the approach's capacity to guide users toward meaningful connections. However, the study acknowledges limitations, including computational overhead in processing large-scale knowledge graphs, the subjectivity of interestingness scoring, and the dependence on LLMs for generating contextually accurate explanations. Despite these challenges, the research demonstrates that integrating LLMs with symbolic reasoning can significantly enhance knowledge discovery and interpretation in cultural heritage datasets.

Pinto et al. [8] propose CURIOCITY, an ontology designed to model cultural heritage knowledge in urban tourism, addressing the limitations of existing ontologies that either represent only partial knowledge or focus solely on indoor cultural heritage, such as museums. The ontology follows a three-tiered architecture—Upper, Middle, and Lower Ontologies—ensuring modularity and adaptability. CURIOCITY is evaluated through a comparative analysis with state-of-the-art ontologies, demonstrating its ability to support automatic population from structured data sources and facilitate the development of tourism-related applications. The results indicate that CURIOCITY effectively integrates indoor and outdoor cultural heritage information, enhancing interoperability and knowledge representation. However, the study acknowledges limitations, including performance bottlenecks in processing large datasets, the need for further refinement in ontology alignment with external knowledge bases, and challenges in ensuring broad adoption across different urban tourism contexts. Despite these constraints, CURIOCITY represents a significant advancement in the semantic modeling of cultural heritage for urban tourism applications.

Carriero et al. [19] introduce ArCo, a knowledge graph for Italian Cultural Heritage, developed using Ontology Design Patterns (ODPs) and the eXtreme Design (XD) methodology to enhance knowledge representation and interoperability. Their study presents a structured approach to modeling cultural heritage data by integrating a network of ontologies that formalize various aspects of cultural properties, including their descriptions, historical contexts, and conservation details. The ArCo knowledge graph, consisting of ∼172.5 million triples, demonstrates significant improvements in semantic querying, interoperability with existing models like CIDOC CRM and Europeana Data Model (EDM), and scalability for diverse cultural heritage applications. The study also introduces TESTaLOD, a tool for unit-testing knowledge graphs, ensuring consistency and validation of ontology-driven cultural data. However, the research identifies challenges, including the complexity of applying ODPs to large-scale cultural heritage data, the evolving nature of ontology requirements, and the need for better tools to automate ontology alignment and testing. Despite these limitations, ArCo represents a major advancement in ontology-driven cultural heritage knowledge representation, offering a reusable and extensible model for digital humanities research and archival institutions.

Tibaut and Guerra de Oliveira [9] propose an improved ontology evaluation framework for cultural heritage information systems, focusing on the semantic interoperability of ontologies. Their study builds upon the OQuaRE evaluation model, introducing three new interoperability metrics—Externes, Composability, and Aggregability—to assess the integration of external ontologies and the modularity of ontology design. The framework is validated through an evaluation of CIDOC CRM (ISO 12217:2014), revealing that while the ontology provides a strong foundation for cultural heritage data exchange, it falls short in certain quality aspects such as structural cohesion, functional adequacy, and reusability. Their findings indicate that integrating external ontologies improves interoperability but requires careful balance to maintain ontology coherence. Limitations of the study include the need for manual refinement in the evaluation process, challenges in automating the interoperability assessment, and the fact that the framework primarily applies to structural ontology evaluation, not domain-specific reasoning tasks. Despite these constraints, the proposed evaluation model enhances ontology assessment methodologies and contributes to improving the reliability and integration of cultural heritage knowledge bases.

Sevilla et al. [20] introduce an ontology-driven framework for the visualization of cultural heritage knowledge graphs, addressing key challenges in the representation and accessibility of digitized heritage data. Their research presents a model that formalizes the visualization of cultural heritage information, incorporating an ontology that integrates existing visualization frameworks such as VUMO and VISO. The study also develops a software framework that enables interactive web-based visualization, tested in projects like SILKNOW and Arxiu Valencià del Disseny. The evaluation results demonstrate that the proposed framework improves knowledge organization and enhances user interaction with cultural heritage datasets. However, the study acknowledges limitations, including challenges in handling uncertainty and granularity variations in cultural heritage data, as well as the difficulty of adapting the visualization model for broader, domain-independent applications. Despite these constraints, the research contributes to the advancement of ontology-based visualization techniques, providing a structured approach to enhancing the accessibility and interpretation of cultural heritage knowledge graphs.

Galera-Rodríguez et al. [21] propose an ontology-driven approach to standardizing the documentation of underground built heritage spaces, a typology that lacks clear classification within traditional architectural or archeological frameworks. Their research develops a methodology based on CIDOC-CRM models to structure and annotate heritage data directly onto three-dimensional models, addressing the challenges posed by the irregular

geometries and difficult accessibility of underground heritage. By integrating web-based tools, such as the Aïoli platform, their approach enhances multidisciplinary collaboration, improving data organization and visualization. The results indicate that this method facilitates the integration of semantic annotation in digital heritage documentation, increasing accessibility and usability for heritage professionals. However, the study acknowledges certain limitations, including the difficulty of ensuring interoperability with broader cultural heritage repositories, challenges in automating certain documentation processes, and constraints in adapting the methodology to diverse heritage contexts. Despite these challenges, their work represents a significant advancement in the semantic modeling and digital preservation of underground built heritage.

Ranjgar et al. [22] provide a comprehensive review of cultural heritage information retrieval, tracing its evolution from early classification systems to modern semantic web technologies. Their study examines how knowledge organization systems, including metadata schemas and ontologies, have shaped cultural heritage data management, with a focus on CIDOC CRM and EDM as key models for interoperability. The research highlights the benefits of ontology-driven approaches in improving data integration, retrieval precision, and machine-readability, enabling more sophisticated querying and knowledge extraction. However, the study also identifies ongoing challenges, such as semantic heterogeneity, the complexity of ontology alignment, and the need for better AI-driven reasoning over cultural heritage data. The authors conclude that while galleries, libraries, archives, and museums have made significant progress in structuring cultural heritage data, there is a pressing need for advanced intelligent applications that leverage existing semantic frameworks for enhanced accessibility and usability.

Ondraszek et al. [23] introduce Viewsari, a knowledge graph designed for digital humanities, specifically modeling Giorgio Vasari's Lives of the Most Eminent Painters, Sculptors, and Architects. Their study applies eXtreme Design (XD) and Ontology Design Patterns (ODPs) to develop a structured representation of Vasari's historical social network, integrating named entity recognition and statistical co-occurrence methods to extract relationships between artists, artworks, and historical contexts. The research highlights the advantages of using agile ontology design methodologies in digital humanities, improving the interoperability and reusability of semantic models. The evaluation demonstrates that Viewsari enhances structured data retrieval and supports scholarly research by enabling the semantic exploration of Renaissance art history. However, the study acknowledges limitations, including challenges in scaling the knowledge graph to incorporate broader art historical datasets, difficulties in maintaining ontological consistency across heterogeneous sources, and the need for continuous stakeholder involvement for refinement. Despite these constraints, Viewsari presents a significant step in advancing ontology-driven knowledge representation for digital humanities research.

Angelis et al. [24] present CHEKG, a collaborative and hybrid methodology for engineering modular and fair domain-specific knowledge graphs. Their research highlights the limitations of existing ontology engineering methodologies in addressing domain-specific requirements such as modularity, agility, and bias mitigation in knowledge graphs. CHEKG introduces a structured approach that combines schema-centric and data-driven methods, integrating human-centered collaboration among knowledge workers, domain experts, and ontology engineers throughout the knowledge graph lifecycle. The methodology is evaluated through its application in semantic trajectory modeling using data from unmanned aerial vehicles, demonstrating improvements in interoperability, data integration, and semantic expressiveness. However, the study acknowledges challenges, including the complexity of coordinating multi-stakeholder contributions, the need for iterative refinements to maintain knowledge graph consistency, and the ongoing difficulty of balancing

automation with expert-driven validation. Despite these limitations, CHEKG represents a significant advancement in ontology-driven knowledge graph development, promoting a modular and transparent approach to domain-specific semantic modeling.

Previous research on ontology-driven cultural heritage frameworks present several limitations that need to be addressed. One key issue is the lack of broad domain coverage in existing ontologies, which often focus on specific conservation cases or heritage types, limiting their applicability to other contexts [15]. Interoperability challenges remain a major concern, as many frameworks struggle to align their ontologies with external knowledge bases and integrate diverse datasets [9,21,25]. The accuracy and reliability of AI-generated knowledge pose another significant challenge, with inconsistencies in outputs, difficulties in validating automatically generated knowledge, and concerns over the ethical implications and bias in AI-assisted ontology engineering [16,25]. Furthermore, performance bottlenecks in querying large-scale knowledge graphs and retrieving complex semantic data highlight the need for optimization in retrieval systems [17,20,25]. Many frameworks also exhibit scalability limitations, making it difficult to expand knowledge models across different heritage domains while maintaining consistency and precision [19]. Additionally, the manual effort required for ontology refinement and validation remains a burden, as automation in knowledge engineering is still reliant on expert oversight to ensure correctness [9,15,25]. Visualization of cultural heritage knowledge also presents granularity and uncertainty issues, making it difficult to create adaptable and user-friendly representation models [20]. Finally, despite advancements in digital heritage technologies, semantic heterogeneity and ontology alignment remain persistent obstacles, limiting seamless integration and cross-institutional collaboration [22].

Our research directly addresses these limitations by developing a domain-specific ontology and knowledge graph for 19th-century Greek bookbinding, ensuring a comprehensive and semantically rich representation of materials, artistic styles, and conservation histories. By integrating LLMs, we enhance ontology expansion while implementing validation mechanisms to ensure accuracy, reliability, and bias mitigation in AI-assisted reasoning. The use of a Neo4j knowledge graph enables efficient querying and advanced pattern recognition, optimizing retrieval performance and addressing scalability concerns. Furthermore, our approach prioritizes interoperability, aligning with established metadata standards such as MARC and Dublin Core to facilitate seamless data integration across institutions. By extending bookbinding documentation beyond textual metadata, our study introduces a more structured, automated, and scalable solution for the study, preservation, and accessibility of historical bookbinding knowledge within the broader cultural heritage landscape.

3. The BookBindKG Framework

This paper introduces BookBindKG, a modular and extensible framework for the construction and evaluation of domain-specific knowledge graphs using LLMs. Designed to model 19th-century Greek bookbinding practices, BookBindKG leverages prompt-engineered interactions with LLMs to generate ontologies, extract structured knowledge, and translate domain queries into executable Cypher code. The framework incorporates a semantically grounded ontology development phase, a rigorous attribute filtering process, programmatic graph construction in Neo4j, and a validation pipeline based on natural language competency questions. Each stage has been crafted to ensure semantic fidelity, data alignment, and interpretability, facilitating its adoption by cultural heritage stakeholders such as conservators, librarians, and historians.

The methods described here reflect a prototyping effort intended to test the viability and precision of LLM-assisted KG construction in a narrowly defined domain. While

technically robust, the workflow is exploratory in nature and is intended as a reusable pattern for future extensions rather than as a final solution.

In what follows, we provide a comprehensive breakdown of the BookBindKG workflow (Figure 1), detailing the methods, technologies, and evaluation metrics used. The proposed approach not only advances the methodological landscape for digital humanities research but also provides a replicable model for applying LLMs in structured knowledge representation tasks.

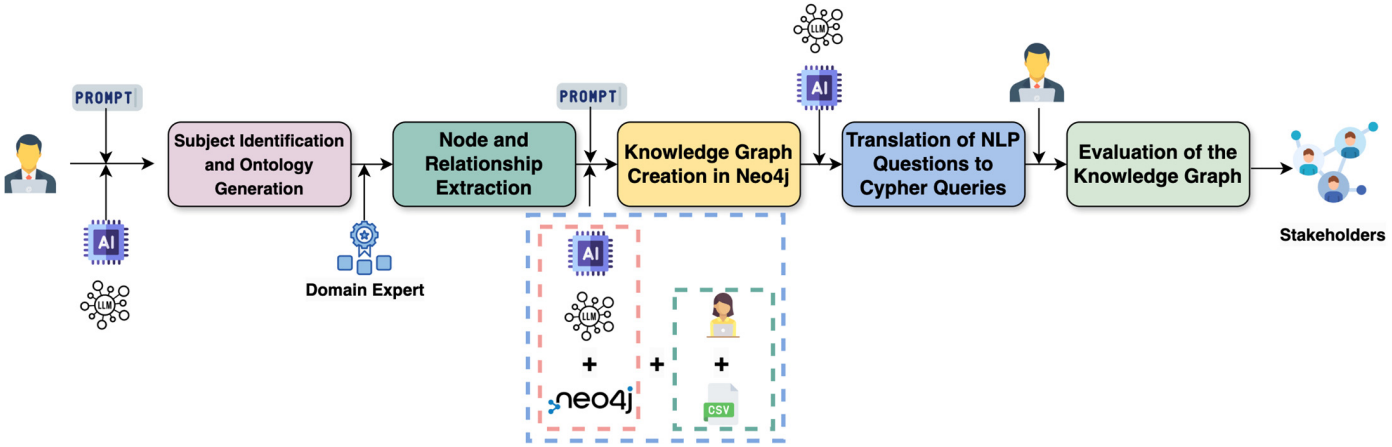

Figure 1. The BookBindKG framework pipeline for constructing and evaluating a domain-specific knowledge graph using LLMs. Each stage integrates LLM-driven prompt engineering, domain expertise, and semantic data modeling. The process begins with subject identification and ontology generation, followed by node and relationship extraction, graph construction in Neo4j, natural language query translation to Cypher, and final evaluation of the knowledge graph's accuracy and completeness. The output is a semantically rich system intended for stakeholder use in the cultural heritage sector.

3.1. Subject Identification and Ontology Generation

The initial phase of the BookBindKG framework involves the formal delineation of the domain to be modeled, specifically 19th-century Greek bookbinding. This stage is crucial in establishing the semantic scope and operational boundaries of the entire knowledge graph development process. To generate a foundational ontology that accurately reflects this domain, a LLM is strategically guided through prompt engineering—a method that constructs a layered, purpose-driven interaction with the model.

The LLM is first assigned a defined role as an ontology engineer. This role-setting ensures that the model approaches the task with an understanding of both methodological constraints and domain-specific expectations. The prompt further instructs the LLM to refrain from generating any output until it has received comprehensive specifications. This controlled interaction enables a stepwise and curated elicitation of knowledge, supporting the generation of a semantically coherent ontology.

Within the defined scope, the ontology is tasked with capturing multiple layers of bookbinding knowledge, ranging from physical descriptors such as size, format, material composition, and structural features, to artistic elements like decorative styles, coloration, and finish. In addition, the ontology must account for metadata pertaining to the cultural and historical context of each artifact, including provenance, authorship, and institutional custody. Conservation and preservation data—such as condition assessments, damage typologies, and treatments applied—also form a critical component of the conceptual model. These modeling choices were directly informed by the limitations of conventional cataloging standards, which lack formal structures for representing the technical components of historical bookbinding. The goal was to extend beyond high-level bibliographic

records and enable a more granular semantic annotation of physical attributes that are vital to scholarly and conservation work.

To further enhance semantic precision, the model is provided with structured metadata from a CSV dataset containing real-world examples of bookbinding documentation. These dataset attributes are mapped against the ontology, with alignment to established metadata standards like Dublin Core. In instances where direct terminological equivalence is lacking, the model is encouraged to semantically interpret the dataset values, aligning them with the most suitable ontological constructs.

Functionality is embedded into the ontology through the formulation of Competency Questions (CQs). These questions represent concrete use cases that the ontology must be capable of addressing. Examples include the following: identifying stylistic and material patterns across different regions or periods; determining the extent and type of damage in preserved volumes; and mapping networks of bookbinders, authors, and sponsors involved in production workflows. These questions serve not only as validation criteria but also as design anchors throughout the modeling process.

Once the ontology has been generated, it is exported in Turtle (.ttl) format, allowing for further refinement and visualization within ontology editors such as Protégé. To enable logical reasoning capabilities, a representative natural language rule is also translated into the Semantic Web Rule Language (SWRL), thereby embedding inferencing logic directly into the ontology's structure. This multi-tiered approach to subject identification and ontology generation ensures that the resulting semantic model is comprehensive, interpretable, and aligned with both the domain's epistemological foundations and practical data realities.

3.2. Node and Relationship Extraction

Following the creation of the ontology, the next phase involves the extraction of structural elements—specifically, the identification of nodes and relationships that will define the knowledge graph's schema. This process requires interpreting the LLM-generated ontology and isolating relevant classes and object properties that align with the available dataset. The aim is to produce a graph structure that is both semantically coherent and directly mappable to empirical data.

To achieve this, ontology classes are initially parsed and assessed for their correspondence to entities or concepts that are represented in the dataset. These include, for instance, classes for individuals (e.g., authors, bookbinders), artifacts (e.g., books, materials), and abstract notions (e.g., preservation conditions, decorative styles). Object properties, which represent relationships between classes, are likewise reviewed for their potential to define meaningful edges between nodes in the graph.

A domain expert plays a critical role at this juncture, functioning as a semantic filter. While the LLM may generate a broad and conceptually rich ontology, not all elements can be operationalized due to the limitations or incompleteness of the structured data. The expert thus curates the ontology by selecting only those entities and relationships for which there is sufficient and reliable data. This ensures that the resulting knowledge graph is grounded in factual representation and avoids speculative or unverifiable assertions. For example, the class ArtisticInfluence—proposed by the LLM to capture stylistic inspiration between bookbinders—was excluded from the implementation. Although semantically meaningful, our dataset lacked explicit references or metadata linking specific bindings to artistic schools or influences. Thus, the absence of verifiable data rendered this class non-operationalizable. In contrast, the class BindingMaterial was fully supported by structured data columns that specified materials such as leather, parchment, or cloth. These were directly mapped to ontology instances and graph nodes, with relationships like HAS_MATERIAL instantiated

accordingly. This contrast illustrates the expert's role in aligning ontology scope with actual data availability, ensuring semantic fidelity without introducing speculative constructs.

Importantly, this step involves an explicit recognition that not all ontology classes will be implemented. The LLM-generated ontology is treated not as a one-to-one blueprint, but rather as a high-level conceptual reference. As such, the ontology serves to guide the modeling strategy and support informed decision-making about entity prioritization and relationship construction.

While the primary data source used for constructing the knowledge graph was a structured dataset containing domain-specific attributes, an LLM-generated ontology was also employed to inform and guide the design of the graph's conceptual model. Although not all class labels and properties in the ontology matched the dataset attributes directly, the ontology served as a high-level semantic reference for identifying core entities, relationships, and modeling patterns within the domain. Where exact correspondences were absent, conceptual alignment was achieved through semantic interpretation, mapping dataset fields to the most relevant or equivalent ontological classes. In this way, the ontology provided a foundational structure that shaped the knowledge graph schema, ensuring consistency with domain knowledge, even when direct one-to-one mappings were not present.

For instance, the ontology included a class labeled BindingOrnamentation, designed to capture aesthetic features such as gilding, embossing, and decorative motifs. However, the dataset did not contain an explicit column labeled as such. Instead, attributes like Decorative Style, Spine Design, and Color Patterns were semantically interpreted to fall under this broader class, and were mapped accordingly. Similarly, the ontology proposed the property UNDERGOES_TREATMENT to represent preservation actions applied to a book. The dataset, however, only included attributes such as Damage Type and Conservation Notes. Through semantic alignment, we associated these attributes with the conceptual intent of UNDERGOES_TREATMENT, allowing the knowledge graph to reflect conservation activity without requiring perfect schema correspondence. These examples illustrate how the ontology functioned as a flexible semantic scaffold, with expert-guided mappings bridging conceptual gaps between ontology design and data structure.

During the ontology review and graph modeling segments, several LLM-generated suggestions were found to be misaligned with the dataset or overly speculative. For instance, the model proposed abstract classes such as SymbolicMeaning or CulturalNarrative, which lacked any direct representation in the source data. Similarly, it occasionally assigned relationships between entities (e.g., between ColorPattern and HistoricalEvent) that implied causality without supporting evidence. A domain expert systematically reviewed these elements, removing constructs not grounded in available metadata or not directly useful for graph instantiation. This expert filtering phase was essential to maintaining semantic accuracy and domain relevance.

3.3. Knowledge Graph Creation in Neo4j

In the third phase of the BookBindKG framework, the ontology-derived schema is transformed into a functional knowledge graph within the Neo4j environment. This process is facilitated through a prompt-engineered interaction with a LLM, which is instructed to assume the role of a Neo4j expert. The model receives a detailed prompt that includes a comprehensive list of column headers from the structured dataset and is tasked with generating a complete Cypher script that operationalizes the ontology into a property graph.

The prompt directs the LLM to distinguish between columns that represent entities (to be modeled as nodes) and those that serve as descriptive attributes (to be represented as

properties). In addition, the LLM must infer semantically coherent relationships among entities based on domain logic. For example, an author is linked to a book via a "WRITES" relationship, while a bookbinder might be associated with the same book through a "BINDS" relationship. The LLM is also instructed to incorporate a LOAD CSV clause into the Cypher script, enabling the structured data to be ingested directly from a local path (file:///BookBindingDataset.csv).

This approach offers several key advantages over traditional ontology-to-graph conversion tools such as the NeoSemantics plugin. While NeoSemantics allows for the importation of RDF/OWL ontologies into Neo4j, it also introduces additional abstraction layers—such as rdf:type, owl:Class, and owl:ObjectProperty—which can obscure the core semantics of the domain. Instead, the manual generation of Cypher scripts ensures precise and transparent control over how nodes and relationships are defined and instantiated. This direct method not only improves interpretability for domain experts but also facilitates optimization for specific use cases.

Moreover, this stage underscores the critical balance between automation and domain-informed customization. While the LLM expedites the generation of graph schema and import code, the structure of the resulting graph reflects curated decisions made during the ontology and node-filtering phases. The resulting model thus remains tightly aligned with domain-specific requirements, while benefiting from the scalability and expressiveness of the property graph paradigm.

By leveraging the LLM to assist in Cypher generation, the framework achieves a semi-automated yet interpretable modeling workflow. This enables rapid prototyping of the graph structure, while preserving a deeper conceptual understanding of the entities and relationships represented. Ultimately, the Neo4j-based knowledge graph resulting from this process is both technically optimized and semantically robust, capable of supporting complex queries and yielding domain-specific insights of high practical relevance.

3.4. Translation of Natural Language Questions to Cypher Queries

In the fourth phase of the BookBindKG framework, the semantic integrity of the constructed knowledge graph is assessed through its ability to respond to domain-relevant queries posed in natural language. These natural language (NL) questions—originally formulated during the ontology development phase—are repurposed as tools for validation and exploration. Rather than functioning as informal prompts, these questions are strategically employed to test the graph's expressivity, logical structure, and alignment with the underlying data.

To operationalize these questions, a large language model (LLM) is engaged once more, now acting as a translator between human language and the Cypher query language used by Neo4j. Prompt engineering ensures that the LLM receives sufficient contextual information about the graph's schema, including node labels, property names, and relationship types. This context enables the model to generate syntactically correct and semantically faithful Cypher queries that can be directly executed against the graph database.

Examples of such NL questions include the following:

- For books housed in Parliament's Library, what are the most common cover materials and binding styles, grouped by subject matter?
- For museum exhibition planning, find books with visually distinctive binding characteristics that would make compelling display items. Include information about their decorative elements, colors, and unique materials.
- Identify trends in conservation treatments over time. What types of treatments were most common for books from different time periods, and how do they correlate with specific deterioration types?

The translation process not only serves as a means of querying the knowledge graph but also provides a mechanism for validating its conceptual soundness. Accurate translation of NL questions into Cypher is contingent on the robustness of the graph schema and its adherence to the ontology-derived structure. As such, this phase bridges the gap between technical modeling and practical utility, demonstrating the graph's capacity to support meaningful, data-driven inquiry in a domain-specific context.

3.5. Evaluation of Knowledge Graph

The final phase of the BookBindKG framework involves the evaluation of the constructed knowledge graph to determine its semantic fidelity, technical robustness, and domain applicability. This evaluation is grounded in the execution of Cypher queries derived from previously formulated natural language questions. Each query tests the ability of the graph to return meaningful and accurate responses reflective of the underlying data and ontological structure.

A domain expert is responsible for assessing the quality of the query results. This human-in-the-loop validation ensures that the interpretation of results accounts for the nuanced knowledge inherent in historical bookbinding practices. The evaluation follows a set of well-established information retrieval metrics [26]:

- Correctness assesses whether the returned answers align with verified truths present in the dataset.
- Precision measures the proportion of retrieved results that are relevant to the original query.
- Recall evaluates how many of the total relevant results were successfully retrieved.
- F1 Score, the harmonic mean of precision and recall, provides a balanced measure of overall performance.

These metrics collectively form a rigorous framework for quality assurance, ensuring that the knowledge graph is not only structurally sound but also functionally effective. By validating the graph against these criteria, the framework establishes its readiness for deployment as a stakeholder-facing knowledge-based system. The evaluation process thus plays a pivotal role in bridging the conceptual model with practical application, confirming that the BookBindKG framework supports reliable, high-fidelity access to domain knowledge.

3.6. Final Outcome

The resulting knowledge graph serves as a semantically rich resource that can be queried, suitable for stakeholders engaged in the study and preservation of historical bookbinding, including conservators, librarians, historians, and cultural heritage professionals. Through its systematic integration of ontology-driven modeling, expert-informed data alignment, and LLM-assisted graph construction, the BookBindKG framework delivers a high-precision knowledge infrastructure capable of supporting both exploratory and analytical tasks.

More importantly, it offers a repeatable and adaptable pipeline, composed of the following sequential stages:

- Ontology Engineering with Large Language Models
- Semantic Filtering and Structured Data Mapping
- Cypher Script Generation for Graph Construction in Neo4j
- Natural Language Question Translation and Evaluation

This modular architecture not only demonstrates methodological rigor but also ensures adaptability to a wide range of cultural heritage and technical domains. Wherever structured metadata and semantic clarity are paramount, the principles underlying Book-

BindKG may be readily applied, fostering broader adoption of knowledge-based systems rooted in both domain expertise and machine-assisted intelligence.

All the script files from KG, cypher queries etc., as well screenshots from every prompting and querying with the results are available in the GitHub (3.15.6 version) repository https://github.com/dimitrisdoumanas19/BookBinders.git (Accessed on 12 April 2025).

4. Results

4.1. Proposed Framework

The ontology engineering process was conducted using a LLM to enhance accuracy and efficiency. We used GPT-4o for the initial phase of ontology creation, based on prior research findings demonstrating [27,28] that GPT-4o delivers the best results in terms of knowledge representation and correct Turtle syntax, minimizing the occurrence of syntactic and semantic errors (Figure 2).

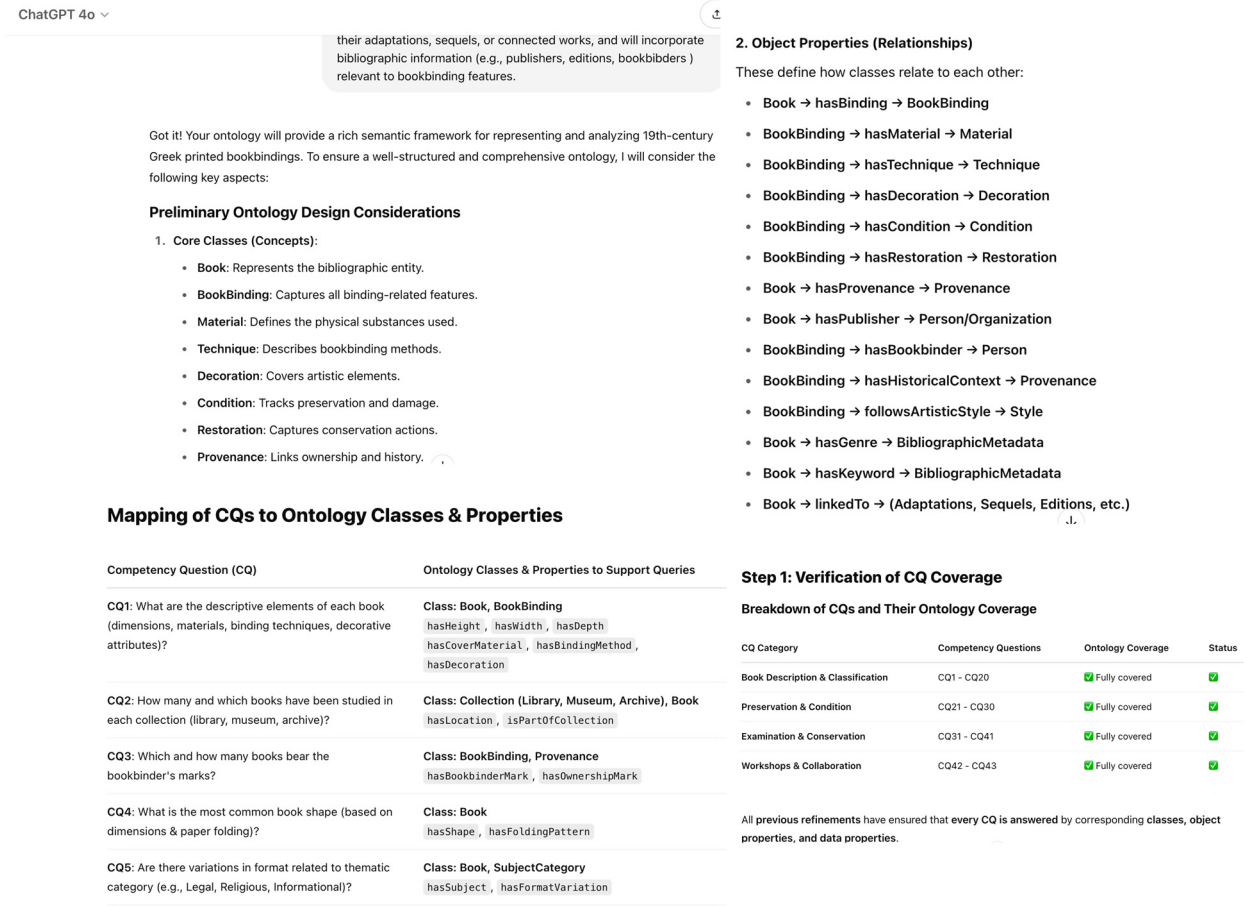

Figure 2. Screenshots from the ontology engineering phase showing the core class definitions, object property relationships, and the mapping of CQs to ontology elements. They also include a verification summary confirming full coverage of all CQs, ensuring the ontology supports detailed reasoning and data integration for 19th-century Greek bookbinding practices.

After developing the initial ontology structure, we continued the process by prompting GPT-4o to extract potential nodes from the generated ontology. To further refine and enhance the extraction and structuring of knowledge, we transitioned to using Claude 3.7 Sonnet (Figure 3). This decision was based on recent studies highlighting that Claude outperforms other LLMs in the specific task of constructing knowledge graphs, showing superior coherence, logical consistency, and node-relationship accuracy [29].

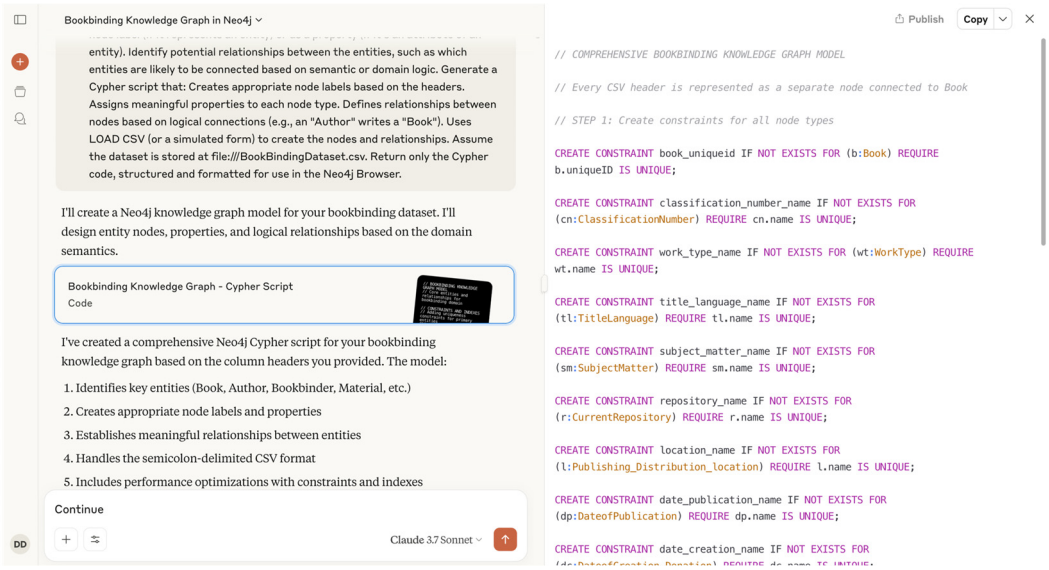

Figure 3. Screenshot from the knowledge graph generation phase, showing a Cypher script used to import and structure bookbinding metadata into Neo4j. The script creates 'Book' nodes from a CSV dataset, assigning key properties and enabling semantic relationships for use in graph-based querying and analysis.

Following the ontology structuring, we proceeded to the graph database implementation phase. We selected Neo4j version 5.24, the latest stable default version available, to ensure full compatibility with knowledge graph operations. A new project named Bookbinders was created with a local DBMS instance. We imported the CSV dataset by placing the file into the designated import folder specified by Neo4j, then accessed the browser interface to initiate the modeling process.

We prompted Neo4j to execute the script for node and relationship creation based on the imported CSV structure. Visualization of the generated knowledge graph was requested directly from the browser. Using the left-side window panel (where the total number of nodes and relationships is displayed), we verified that all intended nodes and relationships were correctly depicted. The completeness of the knowledge graph was thus confirmed.

The following is an example of a prompt used during KG creation:

> *"You are a Neo4j expert. I have a CSV dataset for bookbinding with the following column headers: [Number;UniqueID;ClassificationNumber;. . . . .;reference ID;Citation Details] I want to build a knowledge graph in Neo4j using these headers. Each unique column should be modeled as a node label (if it represents an entity) or as a property (if it's an attribute of an entity). Identify potential relationships between the entities, such as which entities are likely to be connected based on semantic or domain logic. Generate a Cypher script that: Creates appropriate node labels based on the headers. Assigns meaningful properties to each node type. Defines relationships between nodes based on logical connections (e.g., an "Author" writes a "Book"). Uses LOAD CSV (or a simulated form) to create the nodes and relationships. Assume the dataset is stored at file:///BookBindingDataset.csv. Return only the Cypher code, structured and formatted for use in the Neo4j Browser."*

Following the graph construction (Figure 4), we returned to Claude 3.7 Sonnet to assist in translating natural language queries into Cypher queries, Neo4j's native query language. In collaboration with a domain expert from our team, we formulated a set of 10 queries aimed at validating the functionality of the knowledge graph.

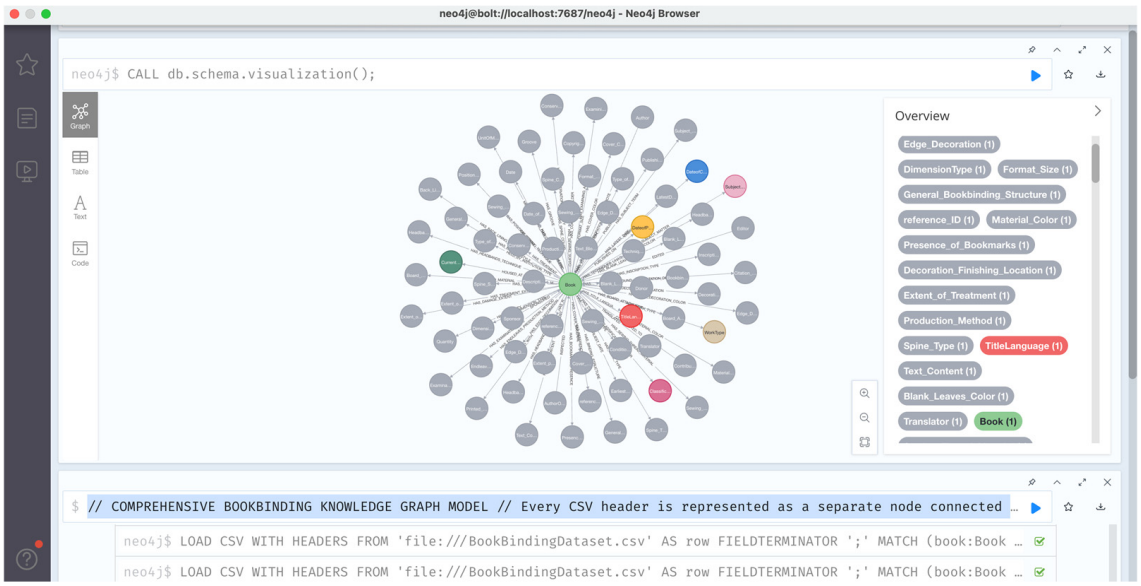

Figure 4. A comprehensive bookbinding knowledge graph model visualized in Neo4j Browser, showing Book as the central node connected to numerous descriptive attributes. The implementation uses Cypher queries to transform structured CSV data into an interconnected graph database.

The intention behind these queries was to replicate realistic questions that would typically arise in the bookbinding domain, with a focus on saving time, reducing manual search effort, and enabling stakeholders to retrieve valuable knowledge efficiently from a complex and interconnected dataset. These queries were compiled with the assistance of domain experts (researchers on book binding and conservation) and were designed to bridge different types of knowledge and entities across the graph, a task traditionally challenging without such automated support:

1. Find all books bound in leather with marbling edge decoration, and show which bookbinders created them, along with their binding structure and style.
2. Identify all books that require urgent conservation, showing their condition type, extent of damage, and deterioration type, sorted by conservation necessity.
3. Which binding techniques and materials were most commonly used during the 18th century? Show the distribution of binding structures, cover materials, and sewing techniques.
4. For books housed in Parliament's Library, what are the most common cover materials and binding styles, grouped by subject matter?
5. Find all books with similar binding characteristics to a specific rare book (UniqueID: '20102301'), showing their current repositories and condition assessments.
6. For conservation planning, identify all books with similar deterioration patterns to help prioritize preservation efforts. Group books by deterioration type, damage extent, and condition to create conservation batches.
7. For museum exhibition planning, find books with visually distinctive binding characteristics that would make compelling display items. Include information about their decorative elements, colors, and unique materials.
8. Which bookbinding materials and techniques show signs of regional variation? Find books with similar binding styles but different geographic origins.
9. Identify trends in conservation treatments over time. What types of treatments were most common for books from different time periods, and how do they correlate with specific deterioration type.

10. Find all books that share similar binding structures but use different combinations of materials throughout their construction. Specifically, identify books that have the same general binding structure, but different combinations of cover materials, board materials, sewing techniques, and spine types. Group these by binding structure to reveal material substitution patterns that bookbinders historically used for the same structural designs.

The following is an example prompt for the first query from natural language to cypher query:

> *"I want you to translate the following queries from natural language to cypher in order to prompt them to a neo4j knowledge graph and get results. the queries are the following: 1. Find all books bound in leather with marbling edge decoration, and show which bookbinders created them, along with their binding structure and style."*

In Figures 5 and 6, examples of the translation and results from querying the KG are presented.

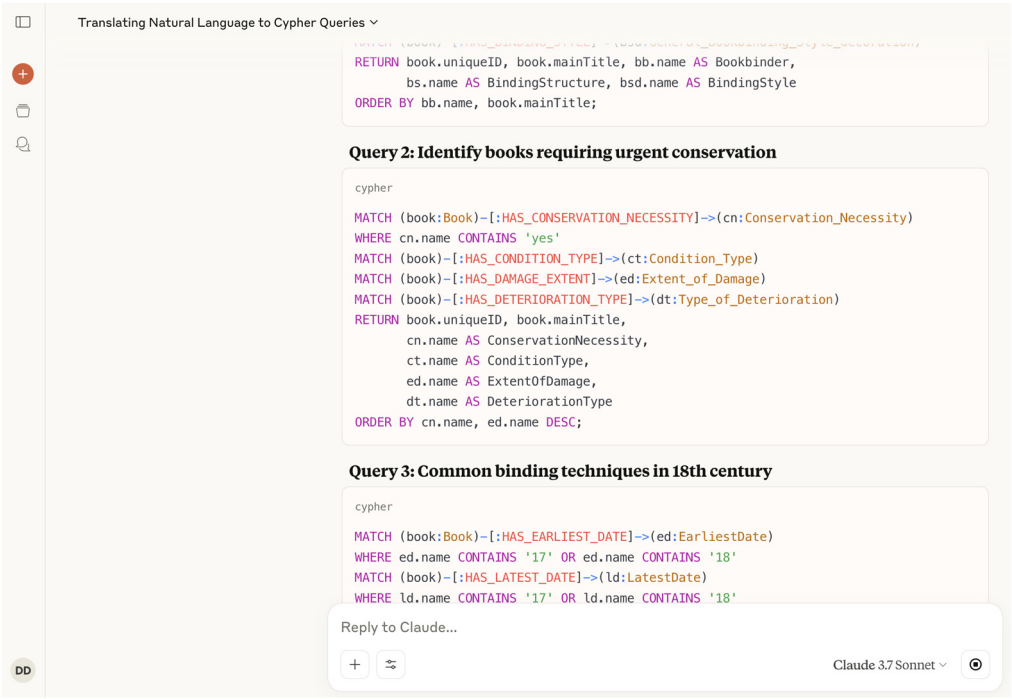

Figure 5. Translation of natural language queries into Cypher using Claude 3.7 Sonnet to enable interaction with the knowledge graph.

All constructed queries were found to be functional, and the result validation phase demonstrated that they returned accurate and complete results. Furthermore, the retrieval time through the knowledge graph interface was at least 20 times faster compared to manual checking methods. The manual checking involved inspecting the CSV file directly (converted into an Excel document and utilizing the functionality provided by Excel, such as sorting, filtering, and searching), which proved significantly slower and prone to oversight.

This performance highlights the potential of such an ontology-driven, knowledge graph-based system to support stakeholders, such as bookbinders, in managing large datasets. It offers a powerful tool for connecting diverse knowledge components, automating retrieval processes, and ultimately facilitating informed decision-making in complex domains.

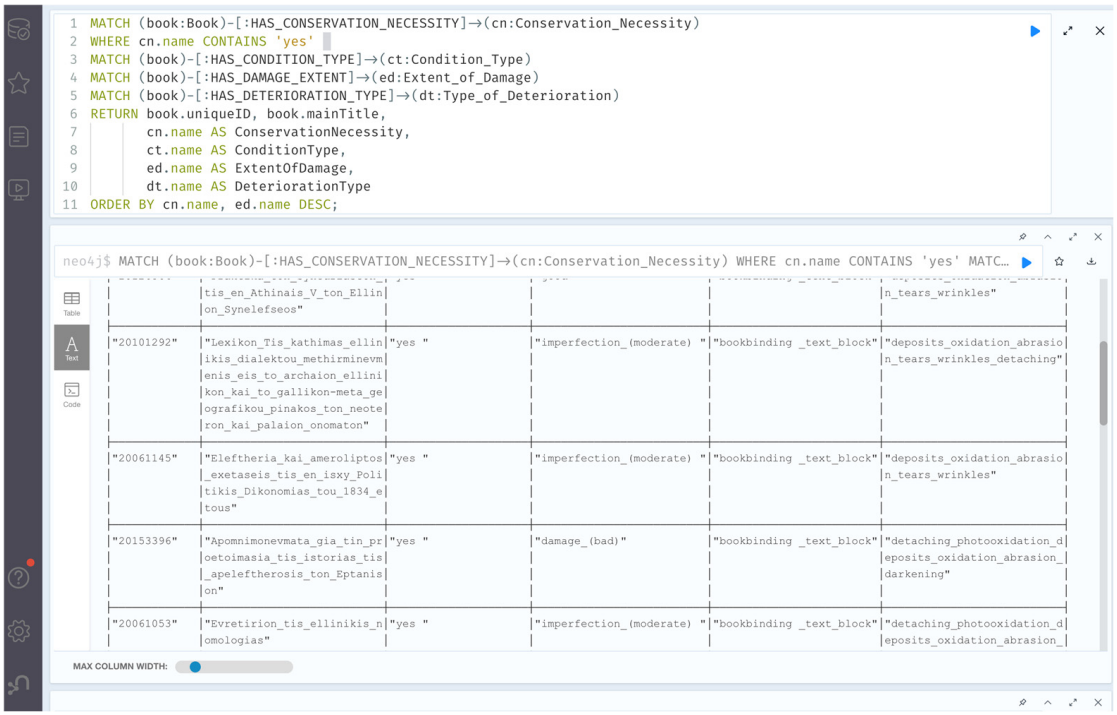

Figure 6. Execution and visualization of Cypher queries in Neo4j, confirming that the knowledge graph structure and relationships function correctly.

4.2. Further Experimentation

While the BookBindKG framework successfully demonstrates an LLM-assisted pipeline for constructing domain-specific knowledge graphs, the true value of semantic technologies lies in their capacity to unify diverse, heterogeneous data sources under a common interpretive lens. As noted in previous research on cultural heritage informatics, isolated datasets—however well-structured—represent only fragments of the broader knowledge ecosystem. The real transformative potential emerges when these knowledge fragments are semantically interconnected with external authoritative sources, enabling queries that transcend the boundaries of any single dataset.

Building upon our established framework, we propose an APOC-based extension that transforms BookBindKG from a self-contained knowledge repository into a dynamic, interconnected semantic hub. This extension leverages Neo4j's APOC library (https://neo4j.com/labs/apoc/ as accessed on 24 April 2025) (Awesome Procedures On Cypher) to establish direct API connections with key semantic web resources, including Wikidata for entity lookup and enrichment, DBpedia for contextual enrichment, and Wikipedia for descriptive content. By implementing these connections, we aim to demonstrate how our domain-specific bookbinding knowledge can be seamlessly integrated into the broader fabric of linked cultural heritage data, enabling more sophisticated queries and insights than would be possible with isolated data alone.

This semantic integration approach addresses a fundamental challenge in cultural heritage documentation: while specialized datasets like our bookbinding collection contain rich domain-specific details, they often lack broader contextual connections that situate artifacts within their historical, artistic, and bibliographic ecosystems. Through this extension, we seek to exemplify how knowledge graphs can serve as unifying interfaces that preserve domain specificity while enabling cross-domain knowledge discovery and enrichment.

In order to validate the knowledge graph's ability to reason over broader literary domains, we expanded our original dataset—which consisted primarily of lesser-known 19th-century Greek book authors—by introducing three renowned figures of modern Greek

literature: Nikos Kazantzakis, Alexandros Papadiamantis, and Odysseas Elytis. For each of these literary icons, we also added one or two of their most influential works—books that are widely recognized as seminal contributions to the evolution of 20th-century Greek literature. This augmentation increased the total number of entries (Book entries) in our dataset to 24.

The rationale behind this expansion was to test the knowledge graph's integration with external semantic resources under conditions where entity linking and biographical enrichment would be more feasible and meaningful. Due to their international reputation, these authors are richly documented across multiple knowledge bases, including Wikidata, DBpedia, and Wikipedia. As such, they provide ideal test cases for evaluating semantic enrichment and cross-domain reasoning. On the contrary, the books in the original dataset were authored by scholars for which information in mainstream knowledge bases is scarce.

Following this update, we prompted Claude with a query designed to test both the internal graph structure and its external linkage capabilities, in order to return the corresponding Cypher query (Figure 7):

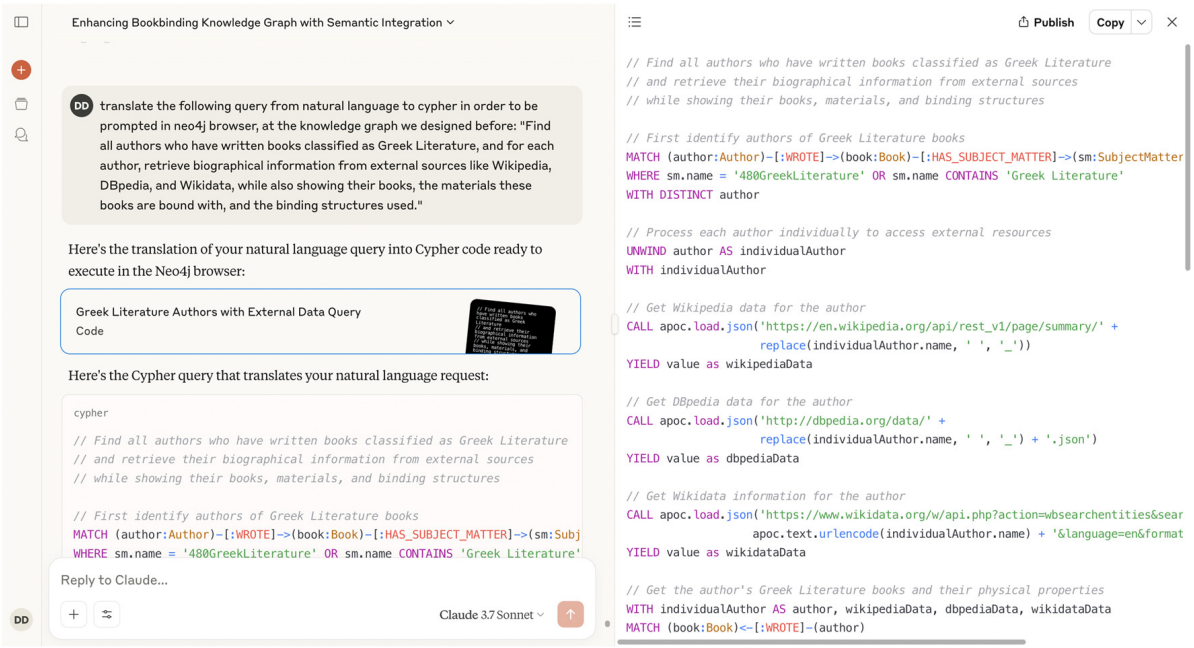

Figure 7. Translation of a complex query into Cypher using Claude 3.7 Sonnet, integrating external semantic data sources (Wikipedia, DBpedia, and Wikidata) into the bookbinding knowledge graph.

> *"Find all authors who have written books classified as Greek Literature, and for each author, retrieve biographical information from external sources like Wikipedia, DBpedia, and Wikidata, while also showing their books, the materials these books are bound with, and the binding structures used."*

The loading of the enriched dataset and the execution of the new Cypher query were then performed using the extended APOC-based semantic integration script described earlier (Figure 8). As shown in Figure 9, the results validated our expectations: Claude successfully identified the famous authors, linked them to external knowledge sources, and returned coherent and richly detailed information about their biographies, their literary contributions, and the physical properties of their books within the knowledge graph.

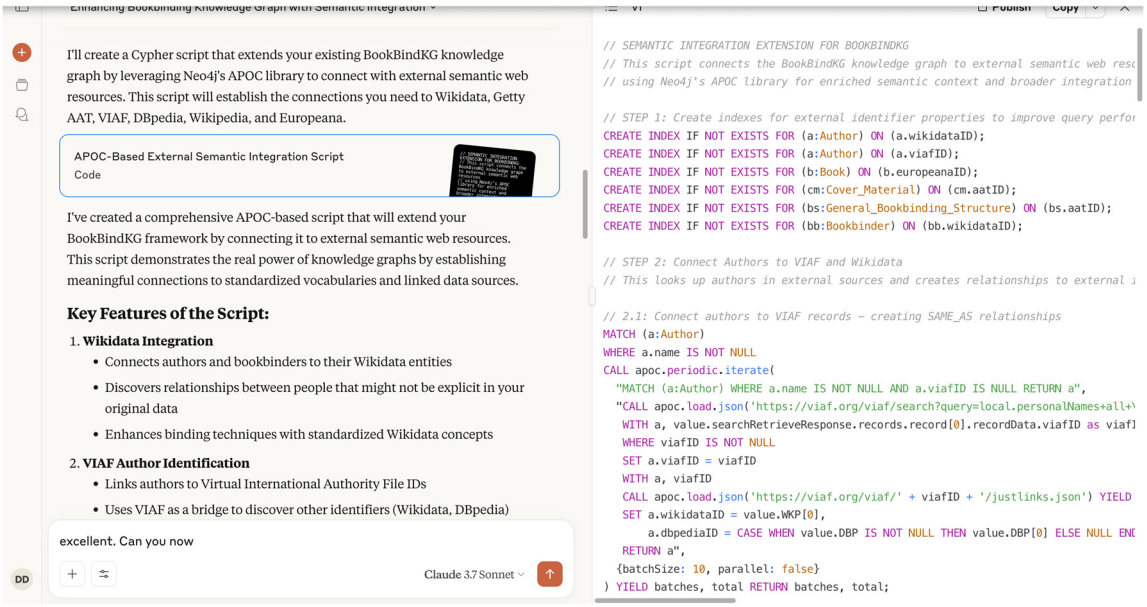

Figure 8. Enhancing BookBindingKG with semantic integration using Wikidata, DBpedia, and Wikipedia articles. The script utilizes Neo4j APOC to enrich knowledge graphs with external linked data.

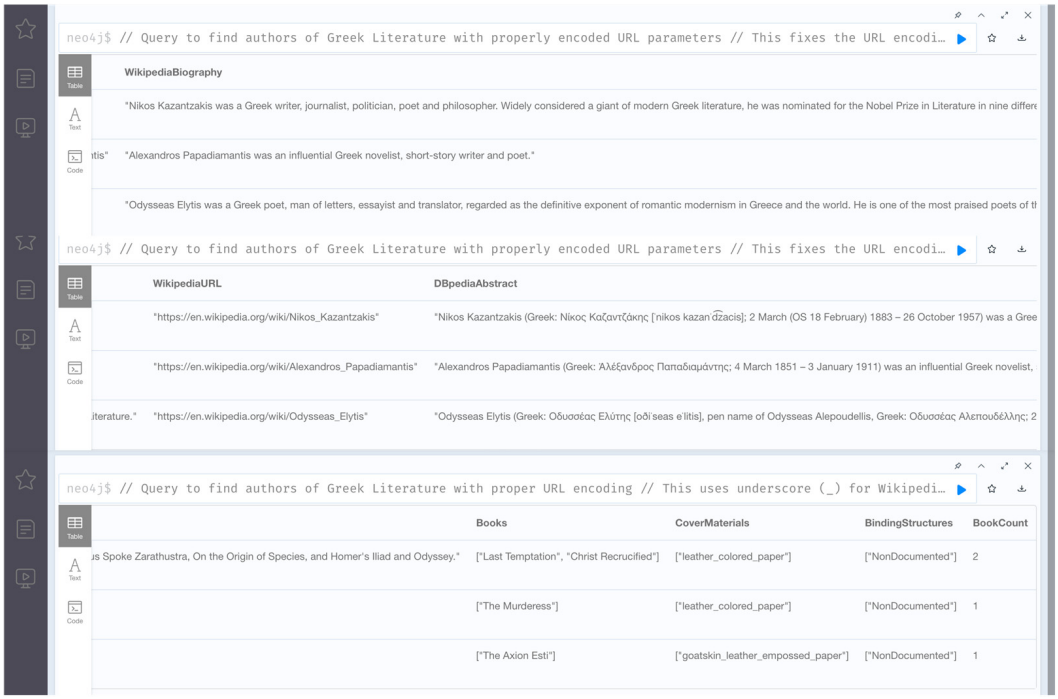

Figure 9. Results from executing a Cypher query that retrieves Greek literature authors along with biographical data from Wikipedia and DBpedia, and links them to their books, cover materials, and binding structures.

5. Discussion

The implementation of the BookBindKG framework relied on a carefully selected suite of tools and technologies that support semantic modeling, scalable graph construction, and accurate data representation. At the core of the process was the use of a state-of-the-art large language model—ChatGPT-4o—which was employed during two key stages: ontology engineering and Cypher script generation for graph instantiation. ChatGPT-4o was chosen

for its advanced contextual reasoning capabilities and its ability to generate structurally coherent, domain-informed outputs from prompt-driven interactions.

For data modeling and visualization, the Neo4j Browser served as the primary graph database interface, with version 5.24 utilized to ensure compatibility with advanced Cypher functionalities and improved performance features. The decision to construct the graph directly within Neo4j—rather than using the NeoSemantics plugin to import the OWL ontology—was deliberate. This choice allowed for fine-grained control over the graph schema, and ensured a direct alignment between data representation and domain requirements. Unlike RDF and OWL imports, which introduce additional abstraction layers such as rdf:type or owl:Class, our approach focused on extracting core entities and relationships from the ontology and explicitly modeling them as nodes and properties. This enhanced the interpretability and usability of the knowledge graph, especially for domain experts unfamiliar with Semantic Web formalisms.

The dataset that underpinned the knowledge graph was originally compiled and structured in Microsoft Excel. It included detailed metadata describing physical, stylistic, historical, and conservation-related aspects of 19th-century Greek bookbinding. Excel was selected for its accessibility and versatility, allowing domain experts to collaboratively curate and refine entries prior to data transformation. This dataset was then mapped to ontology-informed concepts and imported into Neo4j using the LOAD CSV mechanism, with data stored at the local path file:///BookBindingDataset.csv.

Together, these technological choices enabled the seamless integration of human expertise, LLM-driven automation, and graph database capabilities—supporting the creation of a semantically grounded, functionally effective knowledge graph that can serve as a prototype for similar initiatives in cultural heritage informatics.

Beyond the successful technical implementation, our research highlights several key contributions and lessons learned. First, the adoption of an LLM-assisted ontology and knowledge graph construction workflow greatly accelerated the traditionally labor-intensive process of semantic modeling. By integrating human oversight during critical stages such as entity filtering and relationship validation, we achieved a balance between automation and domain-specific accuracy, mitigating risks of semantic drift commonly associated with LLM-generated outputs.

Second, the validation experiments demonstrated that querying over a knowledge graph—even one initially based on a relatively small dataset—can drastically enhance information retrieval efficiency. Compared to manual file inspection, semantic queries across interconnected entities and properties reduced retrieval time by a factor of twenty. This emphasizes the potential of knowledge graphs to support stakeholders such as conservators, librarians, and scholars in rapid, precise decision-making, even in traditionally data-fragmented domains like bookbinding conservation.

Third, the experiment involving the addition of internationally recognized authors (Kazantzakis, Papadiamantis, and Elytis) showcased how the integration of external linked data sources (Wikidata, DBpedia, and Wikipedia) through Neo4j APOC procedures can significantly enrich domain-specific datasets. Through this method, we not only expanded the biographical and contextual metadata available for querying but also validated the extensibility of the BookBindKG framework to cross-domain and cross-institutional knowledge ecosystems.

Moreover, the project underscores the importance of design choices in graph modeling: explicitly defining nodes, properties, and relationships based on the domain ontology—rather than relying on automated OWL imports—led to a more interpretable and user-friendly knowledge graph. This is particularly crucial for cultural heritage pro-

fessionals who may not be trained in Semantic Web formalisms but require accessible, semantically rich tools.

Nevertheless, several challenges were identified during this research. While LLMs demonstrated high performance in generating initial ontologies and Cypher queries, occasional inaccuracies or superficial assumptions required expert validation to ensure semantic precision. Additionally, although external data enrichment via APOC provided powerful capabilities, it also introduced dependency on the completeness and consistency of third-party knowledge bases, which may vary across entities and time.

While the LLMs proved highly effective in accelerating KG and Cypher query generation, they generally performed within the expected boundaries of their design. The outputs did not reflect critical technical errors, but rather required human guidance to ensure strict alignment with the domain-specific context and the structure of the available dataset. For instance, the models occasionally suggested valid Cypher syntax that referenced plausible (but non-existent) graph elements based on general semantic assumptions. These moments were not failures per se, but predictable outcomes given that the LLMs operate without persistent awareness of project-specific schema constraints unless explicitly instructed. In this sense, the role of the human expert was not to correct model flaws but to steer a broadly capable system toward domain-relevant precision. These interactions affirm that LLMs function well as collaborative engines for structured knowledge modeling—efficient, adaptable, and most effective when paired with domain oversight.

Overall, the BookBindKG framework not only offers a scalable and replicable methodology for cultural heritage knowledge graph creation but also lays the groundwork for further developments, such as the following:

- Dynamic enrichment workflows, periodically synchronizing with external linked data sources.
- Expansion into multilingual linked data, supporting broader heritage and literary studies.

By combining machine intelligence, human expertise, and semantic technologies, our work contributes a robust and adaptable blueprint for digital cultural heritage initiatives, demonstrating how structured, interconnected knowledge can breathe new life into historical domains.

6. Conclusions

This study presented BookBindKG, a methodological proof of concept for constructing domain-specific knowledge graphs in the context of historical bookbinding. Rather than aiming to provide a comprehensive solution to cultural heritage metadata challenges, our goal was to explore the current capabilities of LLM-assisted automation when combined with expert-driven data curation. We demonstrated a semi-automated pipeline involving ontology generation, structured data mapping, graph instantiation in Neo4j, and natural language query translation using large language models. The results underscore how human–machine collaboration can accelerate and support the traditionally manual process of semantic knowledge modeling.

A key takeaway from this exploration is the potential of linking structured internal datasets with external semantic resources. By integrating data from sources such as Wikidata, DBpedia, and Wikipedia via Neo4j's APOC library, we showed how even a small domain-specific knowledge graph can be enriched and contextualized within larger knowledge ecosystems. This semantic linking facilitates cross-domain reasoning and highlights promising directions for future metadata augmentation and discovery in cultural heritage.

Our design choices, such as explicitly modeling domain-relevant classes and relationships, and grounding our work in competency questions, aimed to prioritize interpretability, usability, and reproducibility, particularly for non-technical users in the cultural heritage

sector. However, we also acknowledge the limitations of the current work, especially regarding dataset size, real-world validation, and dependency on third-party knowledge bases whose consistency cannot be guaranteed. Furthermore, LLM outputs still require rigorous human oversight to ensure semantic correctness and contextual appropriateness.

Looking ahead, future research should embed similar pipelines into actual institutional workflows and evaluate their impact in collaboration with archivists, conservators, and digital humanists. Further work is also needed to scale the methodology to other heritage collections—such as manuscripts, museum objects, or archival materials—while exploring multilingual integration and more adaptive data enrichment strategies. As a case study, BookBindKG offers a replicable blueprint for further experimentation at the intersection of AI and cultural heritage, contributing to ongoing discourse on how machine learning technologies can responsibly support, rather than replace, domain expertise in heritage knowledge systems. We hope this case study will serve as a practical entry point for professionals in libraries, archives, and museums seeking to understand how emerging AI tools might be integrated into their metadata and knowledge organization workflows.

Author Contributions: Conceptualization, D.D., C.V., M.W. and K.K.; software, D.D.; validation D.D. and K.K.; formal analysis, D.D.; investigation, D.D.; data curation, D.D. and E.N.; writing-original draft, D.D.; writing-review and editing, E.N., C.V., M.W. and K.K.; visualization, D.D.; supervision, C.V., M.W. and K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data are available at the GitHub Repository referred in the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mendoza, M.A.D.; De La Hoz Franco, E.; Gómez, J.E.G. Technologies for the Preservation of Cultural Heritage—A Systematic Review of the Literature. *Sustainability* **2023**, *15*, 1059. [CrossRef]
2. *Handbook of Paper and Board*; Wiley Online Library: Hoboken, NJ, USA, 2006. Available online: https://onlinelibrary.wiley.com/doi/book/10.1002/3527608257 (accessed on 29 April 2025).
3. Diehl, E. *Bookbinding: Its Background and Technique*; Courier Corporation: North Chelmsford, MA, USA, 2013; ISBN 978-0-486-15614-9.
4. Anzalone, J.; Copans, R. Covering the Text: The Object of Bookbinding. *Visible Lang.* **1991**, *25*, 200–240.
5. Roberts, M.T.; Etherington, D. *Bookbinding and the Conservation of Books. A Dictionary of Descriptive Terminology*; Superintendent of Documents; U.S. Government Printing Office: Washington, DC, USA, 1982; ISBN 978-0-8444-0366-3.
6. Tourais, A.; Casanova, C.; Barreira, C.F. Filling the gap: New approaches to medieval bookbinding studies. *J. Mediev. Iber. Stud.* **2022**, *14*, 109–126. [CrossRef]
7. Casillo, M.; De Santo, M.; Mosca, R.; Santaniello, D. Sharing the knowledge: Exploring cultural heritage through an ontology-based platform. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 12317–12327. [CrossRef]
8. Pinto, A.; Cardinale, Y.; Dongo, I.; Ticona-Herrera, R. An Ontology for Modeling Cultural Heritage Knowledge in Urban Tourism. *IEEE Access* **2022**, *10*, 61820–61842. [CrossRef]
9. Tibaut, A.; Guerra De Oliveira, S. A Framework for the Evaluation of the Cultural Heritage Information Ontology. *Appl. Sci.* **2022**, *12*, 795. [CrossRef]
10. Cheng, Y.-J.; Chou, S.-L. Using digital humanity approaches to visualize and evaluate the cultural heritage ontology. *Electron. Libr.* **2022**, *40*, 83–98. [CrossRef]
11. Krötzsch, M.; Thost, V. Ontologies for Knowledge Graphs: Breaking the Rules. In *The Semantic Web—ISWC 2016*; Groth, P., Simperl, E., Gray, A., Sabou, M., Krötzsch, M., Lecue, F., Flöck, F., Gil, Y., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2016; Volume 9981, pp. 376–392. ISBN 978-3-319-46522-7.
12. Chang, Y.; Wang, X.; Wang, J.; Wu, Y.; Yang, L.; Zhu, K.; Chen, H.; Yi, X.; Wang, C.; Wang, Y.; et al. A Survey on Evaluation of Large Language Models. *ACM Trans. Intell. Syst. Technol.* **2024**, *15*, 1–45. [CrossRef]
13. Baldazzi, T.; Bellomarini, L.; Ceri, S.; Colombo, A.; Gentili, A.; Sallinger, E.; Atzeni, P. Explaining Enterprise Knowledge Graphs with Large Language Models and Ontological Reasoning. In Proceedings of the Provenance of Elegance in Computation—Essays Dedicated to Val Tannen (2024), Philadelphia, PA, USA, 24–25 May 2024; Schloss Dagstuhl—Leibniz-Zentrum für Informatik: Wadern, Germany, 2024; pp. 1:1–1:20.

14. Ibrahim, N.; Aboulela, S.; Ibrahim, A.; Kashef, R. A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): Models, evaluation metrics, benchmarks, and challenges. *Discov. Artif. Intell.* **2024**, *4*, 76. [CrossRef]
15. Moraitou, E.; Christodoulou, Y.; Kotis, K.; Caridakis, G. An Ontology-Based Framework for Supporting Decision-Making in Conservation and Restoration Interventions for Cultural Heritage. *J. Comput. Cult. Herit.* **2024**, *17*, 1–24. [CrossRef]
16. Koutsiana, E.; Walker, J.; Nwachukwu, M.; Meroño-Peñuela, A.; Simperl, E. Knowledge Prompting: How Knowledge Engineers Use Large Language Models. *arXiv* **2024**, arXiv:2408.08878. [CrossRef]
17. Linxen, A.; Schmidt, V.-M.; Klinke, H.; Beecks, C. Ontology-driven knowledge base for digital humanities: Restructuring knowledge organization at the library of the Folkwang University of the Arts. In Proceedings of the 2024 IEEE International Conference on Big Data (BigData), Washington, DC, USA, 15–18 December 2024; IEEE: New York, NY, USA, 2024; pp. 2449–2455.
18. Maree, M. Quantifying Relational Exploration in Cultural Heritage Knowledge Graphs with LLMs: A Neuro-Symbolic Approach. *arXiv* **2025**. [CrossRef]
19. Carriero, V.A.; Gangemi, A.; Mancinelli, M.L.; Nuzzolese, A.G.; Presutti, V.; Veninata, C. Pattern-based design applied to cultural heritage knowledge graphs: ArCo: The knowledge graph of Italian Cultural Heritage. *Semant. Web* **2021**, *12*, 313–357. [CrossRef]
20. Sevilla, J.; Samper, J.J.; Fernández, M.; León, A. Ontology and Software Tools for the Formalization of the Visualisation of Cultural Heritage Knowledge Graphs. *Heritage* **2023**, *6*, 4722–4736. [CrossRef]
21. Galera-Rodríguez, A.; Guéna, F.; Algarín-Comino, M. Ontology and collaborative data annotation for the standardisation of documentation processes of underground heritage spaces. *J. Cult. Herit.* **2024**, *68*, 225–236. [CrossRef]
22. Ranjgar, B.; Sadeghi-Niaraki, A.; Shakeri, M.; Rahimi, F.; Choi, S.-M. Cultural Heritage Information Retrieval: Past, Present, and Future Trends. *IEEE Access* **2024**, *12*, 42992–43026. [CrossRef]
23. Ondraszek, S.R.; Petri, G.; Blumenthal, U.; Dieckmann, L.; Posthumus, E.; Sack, H. eXtreme Design (XD) for Ontological Engineering in the Digital Humanities (DH) with Viewsari, a Knowledge Graph (KG) of Giorgio Vasari's the Lives. In Proceedings of the First International Workshop of Semantic Digital Humanities, Co-Located with the Extended Semantic Web Conference 2024 (SemDH), Hersonissos, Zenodo, 27 May 2024. [CrossRef]
24. Angelis, S.; Moraitou, E.; Caridakis, G.; Kotis, K. CHEKG: A collaborative and hybrid methodology for engineering modular and fair domain-specific knowledge graphs. *Knowl. Inf. Syst.* **2024**, *66*, 4899–4925. [CrossRef]
25. Esuli, A.; Sebastiani, F. Evaluating Information Extraction. In *Multilingual and Multimodal Information Access Evaluation*; Agosti, M., Ferro, N., Peters, C., De Rijke, M., Smeaton, A., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6360, pp. 100–111. ISBN 978-3-642-15997-8.
26. Doumanas, D.; Soularidis, A.; Spiliotopoulos, D.; Vassilakis, C.; Kotis, K. Fine-Tuning Large Language Models for Ontology Engineering: A Comparative Analysis of GPT-4 and Mistral. *Appl. Sci.* **2025**, *15*, 2146. [CrossRef]
27. Doumanas, D.; Bouchouras, G.; Soularidis, A.; Kotis, K.; Vouros, G. From Human- to LLM-Centered Collaborative Ontology Engineering. *Appl. Ontol.* **2024**, *19*, 334–367. [CrossRef]
28. Hoppa, J. Claude Converses with Neo4j Via MCP. Graph Database & Analytics. Available online: https://neo4j.com/blog/developer/claude-converses-neo4j-via-mcp/ (accessed on 29 April 2025).
29. Murad, Z. The Power of Knowledge Graphs: Comparing Claude, ChatGPT, and SpaCY in Generating RDFs. Medium. Available online: https://medium.com/@zareenahmurad/the-power-of-knowledge-graphs-comparing-claude-chatgpt-and-spacy-in-generating-rdfs-ef129bdf404a (accessed on 29 April 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.