Journal Title
XX(X):1–18
©The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Modelling Medieval and Renaissance Geographical Knowledge: The IMAGO Knowledge Graph and Semantic Pipeline

Valentina Bartalesi[1] and Nicolò Pratelli[1] and Emanuele Lenzi[1,2]

Abstract

Geographical works from the Middle Ages and Renaissance offer crucial insights into the cultural and intellectual landscapes of their time. However, digital scholarship in this domain remains fragmented, with key historical sources scattered across various physical and digital repositories. The Index Medii Aevi Geographiae Operum (IMAGO), an Italian national research project conducted from 2020 to 2024, addresses this gap by building a semantically enriched, interoperable knowledge graph focused on Latin geographical literature from the 6th to the 15th centuries. By combining expertise in medieval studies, philology, and digital humanities, IMAGO employs Semantic Web technologies and a dedicated ontology extending CIDOC CRM and LRMoo. The project facilitates data integration and reuse by applying Linked Open Data (LOD) principles, thereby enhancing the discoverability and interoperability of cultural heritage data. Beyond the release of the IMAGO knowledge graph, this work contributes a methodological pipeline for semantic modelling, annotation, integration, and publication of data related to medieval and Renaissance geographical works using established Knowledge Representation and Semantic Web standards. The approach is evaluated through a set of scholarly queries. These queries showcase the IMAGO infrastructure's potential for data retrieval and deeper scholarly analysis. Finally, a user-friendly web application further enables access to the knowledge graph via interactive maps, dynamic tables, and exportable formats.

Keywords

Knowledge Representation, Semantic Web, Ontologies, Medieval Manuscripts, CIDOC CRM, Wikidata

Introduction

Geographical works produced during the Middle Ages and Renaissance offer invaluable insights into cultural, intellectual, and scientific paradigms of their respective eras. Indeed, Medieval geographic literature, typically found in encyclopedias and chronicles, provided essential overviews of known territories and served as practical guides for travellers and pilgrims journeying to major religious sites such as the Holy Land, Rome, and Santiago de Compostela. With the advent of Renaissance Humanism, geographical descriptions expanded considerably, driven by the rediscovery and reinterpretation of classical works by Greek authors such as Ptolemy and Strabo. This period also witnessed transformative changes brought about by exploration voyages, which introduced revolutionary accounts of the New World and significantly reshaped contemporary conceptions of physical space.

Despite its importance, digital scholarship in this area remains highly fragmented. Critical historical sources are dispersed across numerous archives, printed editions, and isolated digital repositories, complicating efforts to conduct comprehensive research. Resources such as *Archivio della Latinità Italiana del Medioevo* (ALIM) (Russo 2005), *Archivio digitale della cultura medievale* (MIRABILE) (Pinelli), and *Edizione Nazionale dei testi della Storiografia Umanistica* (ENSU) (Pontari 2006) provide valuable but uncoordinated contributions. To address these challenges, the multidisciplinary project Index Medii Aevi Geographiae Operum (IMAGO) was carried out from 2020 to 2024, funded by the Italian National Research Program (PRIN) (Bartalesi et al. 2022b). The project draws on medieval literature, philology, and digital humanities expertise to develop a comprehensive digital Knowledge Graph (KG) focused on Latin geographical literature from the 6th to the 15th centuries. While the creation of the KG represents a central outcome of the project, the research contribution goes beyond graph publication. The novelty of the work presented in this paper lies in the systematic application of established Knowledge Representation (KR) techniques to a domain - ancient geographical works - that had not previously been addressed using Semantic Web methodologies and technologies, which are able to maximise the interoperability of the collected data. Indeed, the knowledge collected in the graph is formally represented using an ontology, called the IMAGO ontology (Bartalesi et al. 2022a, 2021), which extends several standard vocabularies, such as the ISO standard CIDOC Conceptual Reference Model (CRM) (Doerr 2003), and its extension, the Library Reference Model Object-Oriented (LRMoo)(RIva

[1]Institute of Information Science and Technologies (ISTI), National Research Council of Italy (CNR), Pisa, Italy
[2]Department of Information Engineering (DII), University of Pisa, Pisa, Italy.

Corresponding author:
Valentina Bartalesi, Institute of Information Science and Technologies (ISTI), National Research Council of Italy (CNR), Pisa, Italy.
Email: valentina.bartalesi@isti.cnr.it

et al. 2022). The IMAGO ontology also integrates geospatial standard models such as GeoSPARQL and CRMgeo to accurately represent geographic information. The IMAGO ontology design required the elicitation of domain-specific representational requirements, which were collected in close collaboration with leading scholars in the field of medieval and Renaissance geography and philology. These requirements directly informed the design of the IMAGO ontology, ensuring that the model reflected the research practices and interpretative needs of the target scholarly community. Applying a knowledge representation approach based on the IMAGO ontology provides a robust foundation that also ensures semantic compatibility with other scholarly initiatives and data infrastructures. Indeed, by adhering to Linked Open Data (LOD) principles, IMAGO enhances the discoverability, interoperability, and reusability of historical data in a machine-readable format. Moreover, IMAGO seeks to democratise access to Cultural Heritage (CH) through an intuitive web application that offers interactive maps, dynamic tables, and data export functions (e.g., CSV), thereby supporting academic research and public engagement with medieval and Renaissance geographical knowledge.

The presented work is guided by the following research questions:

1. RQ1: How can heterogeneous knowledge about medieval and Renaissance geographical works be semantically modelled using established KR standards in a coherent and interoperable way?
2. RQ2: How can ontology-driven interfaces enable the population and exploration of a KG by users with different levels of expertise, providing access to complex semantic data through simple interactions and user-friendly representations?
3. RQ3: To what extent does such a semantic modelling and annotation pipeline enable complex scholarly queries that are difficult to perform using traditional methods (e.g. catalogues, critical editions)?

In conclusion, the IMAGO project constitutes a significant advancement in the application of digital humanities methodologies to historical scholarship. By creating a unified, semantically enriched knowledge graph, it addresses longstanding challenges in cultural heritage research and establishes a durable digital infrastructure for future academic inquiry. IMAGO exemplifies how integrating Semantic Web technologies and rigorous scholarly practice can profoundly enhance the study and dissemination of historical knowledge.

Related Work

In recent years, numerous projects in the digital humanities have advanced the development of semantic technologies and LOD frameworks for the study of premodern manuscripts and related cultural heritage materials. Collectively, these initiatives underscore the transformative potential of Semantic Web methodologies in manuscript research by enabling data integration, enhanced provenance tracking, and dynamic scholarly collaboration.

A central example is the Mapping Manuscript Migrations (MMM) project (Hyvönen et al. 2019), which integrates data from three major databases — Bibale, Medieval Manuscripts in Oxford Libraries, and the Schoenberg Database of Manuscripts — through a unified data model based on the CIDOC CRM and FRBRoo ontologies. MMM transforms heterogeneous metadata into a single semantic portal, facilitating large-scale analysis of manuscript histories and allowing researchers to trace provenance across institutional and national boundaries. Closely aligned in focus, the Reconstructing the Phillipps Manuscript Collection project (Burrows 2018) also represents a significant contribution. It centres on the extensive and widely dispersed collection of Sir Thomas Phillipps, evaluating platforms such as Neo4j (Miller 2013) and Nodegoat* for modelling and visualising provenance events. Nodegoat, in particular, enabled in-depth analysis of manuscript dispersal patterns, with special attention to North American institutions (Burrows 2016) and collectors such as Sir Alfred Chester Beatty (Burrows 2017). The resulting models and methodologies have proven valuable for examining the movement and transformation of CH objects. The EU-funded Phillipps project was further developed between 2017 and 2020 by the MMM project, which enhances the technological capabilities of its predecessor (Burrows et al. 2021).

Aligned with MMM's objectives is FactGrid (Faulhaber and Perea Rodríguez 2023), a project of the Gotha Research Centre and the Thuringian State and University Library (ThULB) in Jena. Built on Wikibase and supported by Wikimedia Germany, FactGrid enables users to contribute to and query structured historical data using SPARQL. Its open, collaborative environment supports a wide range of research domains — including manuscript studies — by ensuring all data are Creative Commons-licensed and exportable in multiple formats.

The Digital Manuscripts to Europeana (DM2E) project (Baierer et al. 2017) further illustrates how semantic technologies can enhance manuscript interoperability. DM2E extended the Europeana Data Model (Doerr et al. 2010) to accommodate the specific requirements of manuscript metadata and digitised content across European institutions. Its dataset and scholarly platform provide key infrastructure for studying versioning, provenance, and data reuse, offering insights valuable to developers and researchers.

In North America, the Digital Scriptorium (DS) project (Gilsdorf and Morreale 2024) offers unified digital access to collections of premodern manuscripts. Developed by the Schoenberg Institute for Manuscript Studies, the DS Catalog transforms institutional metadata into LOD enriched with connections to external authorities such as Wikidata. By aggregating manuscript descriptions from a broad network of libraries and museums, DS supports both scholarly inquiry and public engagement, promoting discovery, accessibility, and reuse.

Taking a complementary approach, the China Historical Christian Database (CHCD) (Mayfield et al. 2024) illustrates how LOD frameworks can map historical networks in distinct cultural and temporal contexts. While not

*https://nodegoat.net/

manuscript-focused, CHCD demonstrates how structured data and semantic interfaces can be used to visualise the diffusion of ideas and institutions. Its spatial and relational representations of Christian sites and personnel in China (1550–1950) provide a compelling model for engaging diverse user communities.

Advancements in manuscript structural description are exemplified by the Linked Open Data Based on La Syntaxe du Codex initiative within the Beta maṣāḥǝft project (Liuzzo 2021). This project transforms the codicological model from La Syntaxe du Codex into an ontology implemented through TEI (Ide and Véronis 1995), allowing for more granular structural descriptions of Ethiopian and Eritrean manuscripts. The approach enriches cataloguing practices and highlights the importance of semantic expressiveness in manuscript metadata.

In the African context, the Ontology for Western Saharan Manuscripts (OMOS) (Diakité et al.), developed as part of the BIBLIMOS project (Markhoff et al. 2015), provides a domain-specific framework for the description and analysis of manuscripts from Mauritania and the Western Sahara. OMOS integrates regional cataloguing conventions with international standards such as the CIDOC CRM and FRBRoo, and supports semi-automated enrichment using thesauri like RAMEAU[†], the French national library's thesaurus (Mingam 2015). Its modular, multilingual design addresses the challenges of decentralised manuscript collections often maintained by families or local communities.

Ontologies also enable content-specific searches within manuscript corpora. A notable example is the Medieval Explorations in Neuro-Science (1050–1450): Ontology-Based Keyword Spotting in Manuscript Scans (MENS) project (Alisade et al.), which combines AI-based handwritten text recognition with ontology-driven keyword spotting and named entity recognition. This integration of philological expertise and semantic technologies exemplifies the potential of deep content mining across large-scale digitised manuscript collections.

All these projects reflect an increasing convergence of manuscript studies, Semantic Web technologies, and open data practices. They demonstrate the capacity of LOD not only to standardise and integrate dispersed data but also to enable innovative forms of research into the historical, material, and intellectual dimensions of premodern texts. This convergence around shared conceptual models, particularly the CIDOC CRM, provides fertile ground for initiatives such as the IMAGO project. One of IMAGO's key strengths lies in the interoperability afforded by these standard ontologies: projects that adopt the CIDOC CRM and FRBRoo/LRMoo can relatively easily exchange and integrate data. For example, IMAGO can seamlessly interact with the manuscript graph developed by the MMM project due to their shared conceptual foundations. This interoperability is further enhanced through the use of persistent identifiers from external authorities like Wikidata, allowing scholars to link, for instance, a place mentioned in a travel text with geographic data from a map database or contextual information from a modern encyclopedia. Moreover, semantic models, such as the CIDOC CRM, offer the ability to encode contextual and provenance-related information with a high degree of granularity. Unlike conventional indexes or static maps, ontology-based knowledge graphs can capture not only who authored a text, but also when and where it was written, the sources it drew upon, and even structural features such as the narrative sequence of places in a journey. This semantic richness opens up new research possibilities — for example, enabling complex queries like "find all authors of 14th-century cosmographies who mention Alexandria" - which previously would have required manual collation of disparate sources.

The IMAGO Ontology

A novel contribution of this research lies in the application of KR and Semantic Web technologies to the domain of medieval and Renaissance Latin geographical works - a field that has so far lacked structured and interoperable digital representations. Indeed, existing knowledge is largely dispersed across printed sources and specialised scholarship, hindering systematic analysis and obscuring the historical development of Latin geographical literature. The design and development of the IMAGO ontology is inspired by the well-established METHONTOLOGY framework, which provides a structured and systematic approach to ontology development (Fernández-López et al. 1997). In particular, METHONTOLOGY informed the definition of the main phases of creation, including requirements specification, conceptualisation, implementation, and evaluation. In line with this framework, we adopted an evolving prototype lifecycle that supports continuous refinement and adaptation of the ontology, driven by scholars' feedback. Following the METHONTOLOGY design process, we began addressing the first research question (RQ1). Thus, we collected requirements from domain experts and, based on their analysis, developed a robust conceptualisation of the domain of interest.

The requirement collection was informed by semi-structured interviews with four domain experts actively involved in the IMAGO project, which provided an empirical basis for identifying core entities, relationships, and modelling assumptions. In addition to these interviews, the conceptualisation also drew upon authoritative scholarly works on Medieval and Renaissance Latin geographical manuscripts and printed editions (Pontari et al. 2016; Defilippis 2009; Chiesa 1999).

The materials supporting the conceptualisation are in Italian and are primarily intended for internal use within the project, in close collaboration with domain experts. However, to improve the transparency, we make available two key artefacts that document these first stages of the ontology development process. The first artefact consists of the final version of an annotation template jointly defined with the scholars involved in the IMAGO project (Pontari et al. 2026b,a).This template captures the core conceptual categories and relationships identified as relevant for the domain and represents the outcome of an iterative refinement process based on multiple scholars' interviews

[†] https://data.bnf.fr/en/semanticweb

and discussions. The second artefact is the same annotation template filled in with a concrete example (i.e., Cristoforo Buondelmonti's Descriptio insulae Creta) (Pontari et al. 2026d,c). This example illustrates which values are assigned to each field of the template and clarifies how these values correspond to the IRIs or strings used to instantiate classes in the IMAGO ontology. The resulting set of core categories and relations derived from the requirements analysis serves as the conceptualisation of the IMAGO ontology, capturing the structure and semantics of Latin geographical knowledge in a way that is both theoretically grounded and empirically informed.

Based on the conceptualisation, we implemented the IMAGO ontology[‡]. To maximise its interoperability, the IMAGO ontology extends two foundational vocabularies: the CIDOC CRM (Doerr 2003) and LRMoo (Riva and Žumer 2017) The CIDOC CRM, an ISO standard (ISO 21127:2014), provides an event-based model for representing cultural heritage data, enabling integration with library and archival knowledge. Its event-centric approach allows for flexible, temporal linking of entities, such as connecting a manuscript to its creation event and its printed edition. This structure underpins much of the IMAGO ontology. LRMoo complements this by providing a semantic framework for bibliographic data that supports the integration of museum and library metadata. It enables robust text modelling critical to our project's aims. While IMAGO introduces domain-specific extensions to these ontologies, we intentionally minimised them to ensure interoperability and reduce complexity.

Table 1 presents the alignment between key conceptual categories and IMAGO classes, defined as subclasses of the CIDOC CRM and LRMoo. However, several key conceptual categories are represented using classes directly adopted from established reference ontologies. Specifically, classes such as Work (F2 Expression), Work creation event (F28 Expression Creation), Place (E53 Place), Shelf mark (E42 Identifier), and Date (E52 Time-Span) are used directly from these ontologies without redefining them as new subclasses. This approach was chosen to maximise interoperability with widely adopted cultural heritage standards. In the IMAGO ontology, some of the main properties are defined as subproperties of existing CIDOC CRM properties (see Table 2). Other properties are instead directly adopted from CIDOC CRM, respecting their original domains and ranges. In particular, to model the following relations (R): R(Work creation event,Author), we used the P14 is carried out by; R(Work creation event,Work) we used the R17 created; R(Manuscript,Library) we used the P50 has current keeper; R(Manuscript,Shelf mark) we used the P1 is identified by; R(Manuscript,Folios) we used the P46 is composed of, and R(Manuscript,Date) and R(Printed edition,Date) we used the P4 has time-span.

To facilitate understanding of the ontology structure, Figure 1 presents a diagram of the main classes and properties, along with their relationships. The diagram shows how the ontology reuses and integrates concepts from existing reference models, while introducing the extensions required by our specific use case. This visual representation complements the formal description and

Table 1. The IMAGO ontology classes defined as subclasses of the CIDOC CRM and LRMoo classes.

IMAGO class	Mapping
Author	subclass of E39 Actor
Toponym	subclass of E41 Appellation
Genre	subclass of E55 Type
Manuscript	subclass of F5 Item
Printed Edition	subclass of F3 Manifestation
Library	subclass of F11 Corporate Body
Folios	subclass to E19 Physical Object
Curator/Publisher	subclass of E39 Actor

Table 2. The IMAGO relations and their mapping with the CIDOC CRM properties.

IMAGO property	Mapping
has_genre(Work,Genre)	subproperty of P2 has type
is_identified_by_toponym(Work,Toponym)	subproperty of P1 is identified by
has_curator(Printed edition,Curator)	subproperty of P14 carried out by
has_publisher(Printed edition,Publisher)	subproperty of P14 carried out by

provides an overview of the ontology's main components and their interactions.

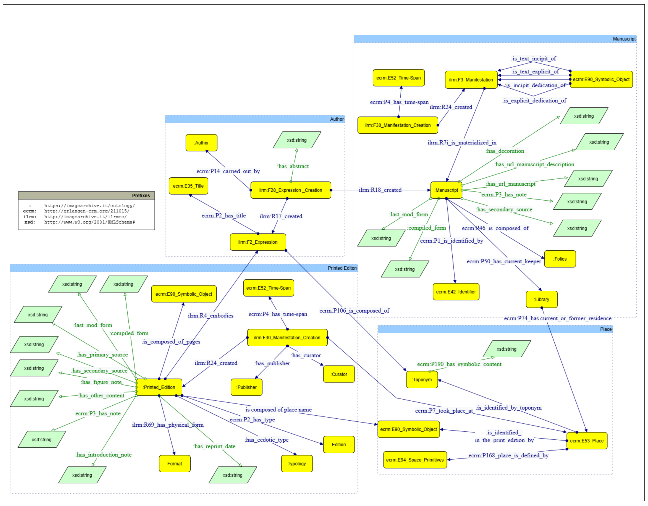

Figure 1. The main classes and principal properties of the IMAGO ontology. Core entities are shown in yellow. Object properties are represented by blue arrows that express relationships between classes. Datatype properties are represented by green arrows, linking entities to literal values displayed in the light green boxes. The larger-framed boxes identify the principal conceptual areas modelled in the ontology, corresponding to Author, Manuscript, Printed Edition, and Place.

Knowledge Graph Population

The IMAGO Corpus

The IMAGO knowledge graph was populated using a corpus curated by the scholars involved in the project. Specifically, they compiled a catalogue of the most important Latin geographical works produced between the 6th and the 15th

[‡] https://imagoarchive.it/ontology/index.html and https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/imago.ttl

centuries. Based on the domain experts' knowledge, the current corpus includes the majority - and certainly the most significant - works within this specialised domain. The catalogue of works and authors can be explored via a public project interface at: https://imagoarchive.it/archive/search.html. For each work in the catalogue, the scholars identified all related manuscripts and printed editions they could trace, including through direct consultation with specific libraries when necessary. Numerical data characterising the corpus, such as the number of authors, works, manuscripts, and printed editions, are summarised in Table 3.

Table 3. Number of authors, works, manuscripts, and printed editions.

Entity type	Count
Authors	237
Works	343
Manuscripts	2333
Printed Editions	431

Then, each manuscript and printed edition was analysed and manually annotated by the experts using the annotation tool described in the following section. The annotation process is entirely expert-driven. No pre-existing structured metadata referring to the works was reused, as such metadata does not exist for this niche corpus. However, when available, a link to a digitalised version of the manuscript or printed edition was recorded and associated with the corresponding resource in the knowledge graph. At this stage, no fully automatic annotation system is adopted, due to the highly specialised and interpretative nature of the sources, which require expert judgement. The annotation tool nonetheless supports the experts through semi-automatic functionalities, such as controlled vocabularies and consistency checks.

An Ontology-Driven Annotation Tool

To address the first part of Research Question 2 (RQ2) - How can ontology-driven interfaces enable the population of a knowledge graph through simple interactions? - we designed and implemented a domain-specific, ontology-driven annotation tool aimed at domain experts. The tool supports the semi-automatic population of the knowledge graph and is built directly upon the representational constraints and structural design of the IMAGO ontology[§]. The annotation tool provides a user-friendly web interface that allows scholars involved in the IMAGO project to collect and annotate data in a structured manner. By tightly aligning the annotation interface with the underlying semantic model, the tool ensures consistency between the conceptualisation phase and data acquisition practices, while enabling non-technical users to contribute high-quality semantic annotations. This design demonstrates how ontology-driven interfaces can effectively guide users in populating a KG while preserving semantic coherence and formal correctness.

In particular, the tool offers the following two main features: (i) Selection of works, authors, libraries, and literary genres from predefined lists provided by the scholars; (ii) Automatic assignment of an IRI to each resource, including authors, manuscripts, literary genres, libraries, places, and their geographic coordinates. Providing predefined lists of works, authors, libraries, and literary genres, the tool helps the scholar save time collecting data. Furthermore, it prevents scholars' errors when entering data manually.

To automatically assign an IRI to each resource, the tool retrieves IRIs from various freely available digital sources, including the Wikidata knowledge base, the MIRABILE digital archive[¶], the Nuovo Soggettario thesaurus (Lucarelli and Bergamin 2013), and the Pleiades gazetteer (Elliott and Gillies 2008).

When IRIs are unavailable in these sources, the tool creates and assigns custom IRIs to works, authors, places, literary genres, and libraries.

For the works and authors, the tool retrieves the IRIs from the Wikidata knowledge base or the Mirabile digital archive. We chose Wikidata for its status as one of the largest general-purpose knowledge bases, offering thousands of entries (Vrandečić and Krötzsch 2014). MIRABILE, on the other hand, is a specialised digital archive focused on the works, authors, and manuscripts that align closely with our project's scope. Based on a list of works and authors provided by the project's scholars, we identified and mapped the corresponding IRIs in both Wikidata and MIRABILE. The mapping process was carried out using a software developed specifically for this purpose [‖]. The software performs SPARQL queries to the Wikidata and MIRABILE endpoints, retrieving a list of candidate IRIs for each work and author. An IMAGO scholar then manually reviews the results, confirming the correct match or selecting the appropriate one when multiple candidate IRIs are found. We were able to automatically retrieve 98% of the IRIs for the works and 96% for the authors.

For the libraries, the tool retrieves the IRIs from Wikidata, which contains thousands of descriptions of geographic entities and libraries. Starting from a list of libraries provided by the scholars, we performed SPARQL queries to the Wikidata endpoint to retrieve the corresponding IRIs. In addition to the IRIs, we extracted the English name of each library and the IRI of the place where it is located. For each place, we also retrieved the English name, geographic coordinates (latitude and longitude), and the country to which it belongs. These two pieces of knowledge are used to visualise the places on a map. We were able to retrieve about 41% of the library IRIs in Wikidata.

For the literary genres, the tool retrieves the IRIs from the Nuovo Soggettario, a thesaurus created and maintained by the National Central Library of Florence and developed in compliance with the International Federation of Library Associations and Institutions (IFLA) (IFLA 2023). Based on a list of 40 literary genres provided by the scholars (e.g., Latin historiography and Latin dramaturgy), we successfully

[§] https://github.com/AIMH-DHgroup/IMAGO/tree/main/tools/annotation-tool
[¶] http://www.mirabileweb.it/
[‖] https://github.com/AIMH-DHgroup/IMAGO/tree/main/tools/linking

mapped approximately 95% of them to corresponding entries in the Nuovo Soggettario thesaurus.

Finally, as one of the main aims of the IMAGO project is to create an index of toponyms cited in the works, the tool retrieves the IRIs of those toponyms from external linked data sources. To achieve this goal, when a scholar enters a toponym into the tool's interface, a SPARQL query is sent both to the Wikidata endpoint and to the Pleiades gazetteer - an open-access digital gazetteer for ancient places - to retrieve the corresponding IRIs. In particular, the use of Pleiades enables us to handle variant name forms, especially those found in medieval works.

Knowledge Graph Implementation

Once the data related to the works have been fully entered by the scholars using the semi-automatic tool, an OWL 2 DL knowledge graph is automatically generated. The knowledge graph is created by a Java-based triplifier software (Pratelli et al. 2026f) developed by the authors, which models the data according to the IMAGO Ontology. The IMAGO Ontology imports two external reference ontologies: the official OWL implementation of CIDOC CRM (CIDOC CRM v7.1.1, Erlangen version (of Erlangen-Nuremberg; Germanisches Nationalmuseum; Zoologisches Forschungsmuseum Koenig 2013)), and an OWL version of LRMoo developed by the authors, since no official OWL implementation of LRMoo v0.7[**] was available. In addition, the knowledge graph imports a thesaurus developed by the authors to formally represent literary genres [††]. The resulting knowledge graph is stored in an Apache Jena Fuseki triplestore (Jena 2014), which provides a public SPARQL endpoint (Pratelli et al. 2024a) for querying and exploring the data. The SPARQL endpoint exposes three named graphs, each serving a distinct purpose. The main graph (Pratelli et al. 2026b) contains the knowledge graph of the works, modelled according to the IMAGO ontology. This graph contains 108,214 triples. A second graph contains only the toponyms (Pratelli et al. 2026e). These toponyms are modelled using the IMAGO ontology, resulting in 2,709 triples. We decided to separate toponyms into a dedicated graph because we are currently conducting a study of the formal representation of place names as attested in ancient manuscripts. However, this research activity falls outside the scope of the present paper and is mentioned here only to motivate the choice of modelling toponyms in a separate graph, which facilitates their independent analysis and future extension while maintaining consistency with the overall knowledge graph. Finally, a third graph includes a portion of the Mapping Manuscript Migrations (MMM) knowledge graph (Pratelli et al. 2026a). A preliminary integration experiment with this graph is currently underway, as mentioned in the Discussion section, but it is outside the scope of the present paper. This graph contains 7,932 triples. Overall, the three named graphs collectively contain 118,855 triples. However, for the purposes of the research presented in this paper, only the main knowledge graph and the toponym graph are considered, excluding the MMM graph, resulting in a total of 110,923 triples.

Numerical data characterising the resources included in the IMAGO main knowledge graph are partially summarised in Table 3, which reports the number of works, authors, manuscripts, and printed editions catalogued by the scholars involved in the IMAGO project. In addition to these entities, the main knowledge graph also incorporates 417 libraries and 652 places, all of which were annotated using the dedicated annotation tool. Furthermore, the toponym graph includes 385 toponyms related to the annotated places.

To guarantee the logical consistency of the knowledge graph, we employed the open-source OWL 2 DL reasoner Openllet (Openllet 2023). The reasoner was used as a validation tool and applied to four main tasks. First, it ensured the global logical consistency of the knowledge graph by verifying that no logical contradictions arise from combining classes, properties, and axioms. Second, Openllet was used to validate the class hierarchy, checking that all inferred class memberships and subclass relations conform to the structure defined by the IMAGO ontology. Third, the reasoner was applied to data integrity checking, verifying that individuals comply with the declared domain and range constraints of properties and that no violations of ontology axioms occur. Finally, Openllet was used to assess whether the knowledge graph supports SPARQL queries, ensuring that the inferred knowledge can be correctly derived through reasoning. In particular, these queries can automatically retrieve inferred superclass, superproperty, and inverse property relationships (Pratelli et al. 2026d).

Openllet successfully passed all these validation tasks for the knowledge graph. To ensure transparency and reproducibility, we have made available in a public GitHub repository both the input knowledge graph used in the validation process and the corresponding output knowledge graph produced after reasoning, including the inferred triples generated by Openllet (Pratelli et al. 2026c).

Knowledge Extraction and Evaluation via SPARQL Queries

To address Research Question 3 (RQ3) - To what extent does such a semantic modelling and annotation pipeline enable complex scholarly queries that are difficult to perform using traditional methods (e.g., catalogues, critical editions)? - we defined and implemented a set of queries over the IMAGO knowledge graph. These queries were designed to evaluate the knowledge graph's capacity to support both scholarly research and public engagement by facilitating the analysis and exploration of complex data. To this end, we gathered a representative set of research questions formulated by scholars involved in the IMAGO project. These questions were identified during plenary meetings with the IMAGO community and through an analysis of the primary research objectives of participating experts, who often focus on the critical edition of ancient texts. Within the context of the IMAGO project, a lemma is defined as a pair consisting of an author and the title of a work. Key queries include, for example, retrieving lemmas by author or title, identifying all works that reference a specific geographic location,

[**] This is the latest version released at the time the IMAGO ontology was created: https://cidoc-crm.org/lrmoo/ModelVersion/lrmoo-f.k.a.-frbroo-v.0.7.

[††] https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/imports/TGG20220221.owl

and finding manuscripts preserved in particular libraries. By implementing these queries, we demonstrate how the semantic modelling enables complex explorations that would be difficult or cumbersome using traditional resources.

We identify six types of knowledge extraction targets. These targets represent distinct yet complementary search functionalities, implemented through specific SPARQL queries. Each query addresses a particular dimension of medieval and Renaissance geographical works, enabling a nuanced exploration of the data. Specifically, we extracted information related to:

1. Lemmas based on the author's name or the title of the work.
2. Lemmas categorised by literary genre.
3. Works in which a specific place is mentioned.
4. Manuscripts held in a particular library.
5. Printed editions published in a given location.
6. Manuscripts dated to one specific century.

The information extracted by these queries overall covered the interests of the IMAGO scholars. To further validate the relevance of these queries, we also involved users and experts who were not part of the design process. The final evaluation was conducted with a group of scholars specialised in philology applied to the study of medieval and Renaissance geographical works during the concluding international IMAGO conference[‡‡]. This event brought together internationally recognised researchers in the field, and a dedicated slot was reserved for discussing the queries defined in the project. All participating scholars agreed that these queries were highly relevant for the study of this discipline and successfully retrieved information meaningful for historical and philological research. Moreover, the evaluation highlighted that the queries effectively capture the types of questions scholars are likely to pose in real research scenarios, including cross-referencing authors, places mentioned in the works, manuscripts and printed editions of a single work. The experts emphasised that the structured nature of the knowledge graph, together with these well-defined queries, enables more efficient exploration and comparison of sources than would be possible using traditional catalogues or critical editions alone. This external validation provides evidence that the ontology-driven knowledge graph and its associated queries not only reflect domain knowledge accurately but also support practical, high-level scholarly investigations across multiple dimensions of medieval and Renaissance geographical literature.

Discovering New Knowledge To demonstrate that the knowledge graph effectively supports data analysis and contributes to the discovery of new insights, we implemented one SPARQL query for each of the six knowledge-extraction requirements outlined in the previous section. In particular, for each requirement, we implemented a specific query which satisfies it.

The code of the queries is available in Appendix A. The SPARQL endpoint used to run these queries is accessible at `https://imagoarchive.it/sparql/`.

As a notational convention, in what follows, we use prefixed names (e.g., `ecrm:`) to denote terms re-used from other vocabularies. The prefix `xsd:` refers to the XML Schema Definition namespace, which provides standard data types such as strings, dates, and integers. The `rdf:` and `rdfs:` prefixes correspond to the RDF and RDF Schema namespaces, respectively, and are fundamental for expressing triples and defining classes, properties, and relationships in RDF. The `ecrm:` prefix designates the Erlangen implementation of the CRM. The `ilrm:` prefix refers to the IMAGO project implementation of LRMoo ontology, an extension of CRM tailored to bibliographic and library contexts. Finally, the default prefix `:` denotes terms of the IMAGO ontology (e.g., `:Author`).

Q1: Retrieving of the lemmas based on the author's name and work title. The SPARQL query retrieves lemmas based on the author's name or work title. Specifically, it identifies instances of the `ilrm:F28_Expression_Creation` class, which represents events of intellectual production. Each such event links an author (typed as `:Author`) to a created work (typed as `ilrm:F2_Expression`) through the property `ilrm:R17_created`.

Authors are identified via the `ecrm:P1_is_identified_by` property, from which both the primary name (`ecrm:P190_has_symbolic_content`) and any aliases (`:has_alias`) are extracted. The title of the created work is similarly retrieved through the property `ecrm:P102_has_title`.

The query uses exact string matching through two `FILTER` clauses: one to match the work title (`"Descriptio insule Crete"`), and another to match the author's name (`"Christophorus Bondelmontius"`). The query returns the creation event, the title of the work, the author's canonical name, and any associated aliases. Results are ordered alphabetically by the author's name and the title of the work.

Although the current query uses exact matches, it can be extended to support dynamic or partial searches via case-insensitive regex filters or parameterized input. In such a case, frontend search forms could inject user-provided values into the query at runtime to enable broader retrieval, for example by matching substrings or name variants.

Table 4. Results of the query in Figure .

title	authorName	alias
Descriptio insule Crete	Christophorus Bondelmontius	Cristoforo Buondelmonti, Cristoforo de' Buondelmonti, Cristoforo dei Buondelmonti, Cristoforo Bondelmonti, Cristophoro Buondelmonti, Cristoforo Bondelmonti

Q2. Retrieving all works categorised under the literary genre of Cartography. This SPARQL query is designed to retrieve all works categorised under the literary genre of *Cartography*, which is identified in the IMAGO thesaurus by the IRI `https://imagoarchive.it/thes/tid/10003`. For each work, the query also extracts the title and the name of the associated author.

[‡‡] `https://imagoarchive.it/imago-convegno-2024-03-19.html`

The query begins by identifying instances of the class ilrm:F28_Expression_Creation, which represents intellectual creation events. Each event is linked to a created work (ilrm:F2_Expression) via the property ilrm:R17_created. The individual responsible for the creation is retrieved using the property ecrm:P14_carried_out_by and must be an instance of the class :Author.

To obtain the author's name, the query uses the property ecrm:P1_is_identified_by, resolving to the textual content representing the name. The work's title is retrieved similarly via the property ecrm:P102_has_title.

The query includes a filter to return only those works explicitly associated with the *Cartography* literary genre, achieved through the custom property :has_genre and the specified genre IRI. The results are sorted alphabetically by author name to facilitate reference and comparison.

The output of this query is presented in Table 5.

Table 5. List of works classified under the genre of Cartography, along with their authors.

Author	Work Title
Christophorus Bondelmontius	Descriptio insule Crete
Christophorus Bondelmontius	Liber insularum archipelagi
Franciscus Petrarcha	Itineriarum breve de Ianua usque ad Ierusalem et Terram Sanctam
Hugo de Sancto Victore	Descriptio mappe mundi
Leo Baptista Albertus	Descriptio urbis Romae

Q3. Retrieving all works in which the toponym Genoa is explicitly mentioned. This SPARQL query retrieves all works in which the Italian city of *Genoa* is explicitly mentioned. In the IMAGO knowledge base, Genoa is represented by the IRI https://imagoarchive.it/ontology/resources/toponym/genova. For each identified work, the query also extracts the work's title and the name of its author.

The initial structure of the query follows the same pattern as previous examples: it begins by identifying instances of the class ilrm:F28_Expression_Creation, which models intellectual creation events. These events are linked to the resulting work via ilrm:R17_created, and to the responsible agent (author) via ecrm:P14_carried_out_by. The author's name is obtained by traversing the properties ecrm:P1_is_identified_by and ecrm:P190_has_symbolic_content, while the title of the work is retrieved via ecrm:P102_has_title and the same symbolic content property.

To restrict results to only those works that mention Genoa, the query uses the property ecrm:P106_is_composed_of, which associates a work with a referenced toponym. The toponym is explicitly bound to the Genoa IRI using the operator BIND. Additional geographic context is obtained by linking the toponym to a ecrm:E53_Place instance using the custom property :is_identified_by_toponym. From this place instance, geographic coordinates can be accessed via the property ecrm:P168_place_is_defined_by.

The results of this query are presented in Table 6.

Table 6. List of works where the toponym *Genoa* is mentioned, including authors and titles.

Author	Work Title
Franciscus Petrarcha	Itineriarum breve de Ianua usque ad Ierusalem et Terram Sanctam
Iacobus Bracellus	Descriptio orae Ligusticae
Opus sine auctore	Itinerarium de Brugis
Antonius Gallus	Commentarius de Genuensium maritima classe in Barchinonenses expedita
Antonius Gallus	Commentarii rerum Genuensium
Ludolphus de Sudheim	De itinere Terrae Sanctae
Antonius Gallus	De navigatione Columbi per inaccessum antea Oceanum comentariolum
Iohannes Porta de Annoniaco	Itinerarium cardinalis

Q4. Retrieving all works whose physical manuscripts are currently preserved at the National Central Library of Florence. This SPARQL query retrieves all works whose physical manuscripts are currently preserved at the National Central Library of Florence, which is identified by the IRI http://www.wikidata.org/entity/Q460907. Specifically, the query extracts the manuscript identifier, shelf mark, number of folios, the name of the holding library, and the name of the associated geographic location.

As in previous examples, the query begins by identifying intellectual creation events (ilrm:F28_Expression_Creation), which are linked to the corresponding works (ilrm:F2_Expression) via the property ilrm:R17_created. The responsible author is retrieved through the property ecrm:P14_carried_out_by, and their name is obtained as usual through the property ecrm:P1_is_identified_by. The title of the work is accessed in a similar manner via ecrm:P102_has_title.

To identify the physical manuscripts, the property ilrm:R18_created connects the intellectual creation to a manuscript, and ilrm:R7i_is_materialized_in links the manuscript to its physical manifestation. The manuscript has to be currently held by the specified library, as indicated by the property ecrm:P50_has_current_keeper. Its shelfmark is retrieved via the path ecrm:P1_is_identified_by, and the number of folios is obtained using ecrm:P46_is_composed_of.

To provide geographic context, the query traverses ecrm:P74_has_current_or_former_residence to locate the place associated with the library. This place is linked to a toponym via the custom property :is_identified_by_toponym, from which both the place name and geographic coordinates are extracted using ecrm:P190_has_symbolic_content and ecrm:P168_place_is_defined_by, respectively.

The results are presented in Table 7, and sorted by place name, library name, and manuscript shelf mark.

Table 7. List of manuscripts preserved at the National Central Library of Florence.

Place	Library	Shelf mark and Folios
Firenze	Biblioteca Nazionale Centrale	Banco rari 50 , 123v-124r
Firenze	Biblioteca Nazionale Centrale	Conv. Sopp. C.VII. 1170, 1r-69r
Firenze	Biblioteca Nazionale Centrale	Conv. Soppr. C.8.2861, 1r-27r
Firenze	Biblioteca Nazionale Centrale	Conv. soppr. F.4.733, 29r-43r
Firenze	Biblioteca Nazionale Centrale	Conv. soppr. I.2.37, 92v-156v
Firenze	Biblioteca Nazionale Centrale	Conventi Soppressi J I 45, 1r-83v
Firenze	Biblioteca Nazionale Centrale	Fondo Conventi Soppressi, J IV 2 (cod. 283), 37r-38v
Firenze	Biblioteca Nazionale Centrale	II, II, 312, 2r-50v
Firenze	Biblioteca Nazionale Centrale	Landau Finaly 5
Firenze	Biblioteca Nazionale Centrale	Landau Finaly 58, 1r-182v
Firenze	Biblioteca Nazionale Centrale	Magl. XXII.22, 107r-121v
Firenze	Biblioteca Nazionale Centrale	Magliabecchiano XIII 7, 1r-47r
Firenze	Biblioteca Nazionale Centrale	Magliabechiano XIII, 38, 1r-194r

Q5. Retrieving all printed editions that were published in the city of Bologna. This SPARQL query retrieves all printed editions of manuscripts that were published in the city of *Bologna*, Italy, identified in the IMAGO knowledge base by the IRI `http://www.wikidata.org/entity/Q1891`. In addition to the core bibliographic data, the query extracts contextual information such as publication place, publisher, time-span, and curator, where available.

The query begins by identifying intellectual creation events (`ilrm:F28_Expression_Creation`) and the associated works (`ilrm:F2_Expression`), along with the name of the author. Author names are obtained by traversing the property path `ecrm:P1_is_identified_by` followed by `ecrm:P190_has_symbolic_content`, and the titles of the original manuscripts are retrieved similarly via `ecrm:P102_has_title`.

What distinguishes this query is its focus on printed editions (`?print_edition`) that *embody* the original work, linked using the property `ilrm:R4_embodies`. Each printed edition has its own associated creation event, identified via `ilrm:R24_created`, which reflects a separate act of intellectual or material production.

To restrict the results to those editions published in Bologna, the query uses the property `ecrm:P7_took_place_at`, with the toponym bound to the IRI for Bologna. This toponym is further resolved into a readable place name using the custom property `:is_identified_by_toponym` and `ecrm:P190_has_symbolic_content`.

Additional metadata is retrieved through optional clauses:

- The time-span of the printed edition's creation (`ecrm:P4_has_time-span`)
- The publisher (`:has_publisher`)
- The curator involved in the printing process (`:has_curator`)

All optional elements follow standard CIDOC CRM property chains.

The results of this query are shown in Table 8.

Table 8. Printed editions produced in Bologna (Italy)

Place	Publisher	Date
Bologna	Deputazione di Storia Patria	1990
Bologna	Tipografia Alfonso Garagnani e figli	1896
Bologna	Domenico de' Lapi	1462
Bologna	Azzoguidi	1917

Q6. Retrieving metadata about manuscripts produced within the 10th and 11th centuries. This SPARQL query is designed to retrieve metadata about manuscripts produced within the 10th and 11th centuries. The query targets resources classified as manuscripts and filters them based on their associated time spans to ensure temporal relevance to the defined historical period (901–1100 CE).

The query begins by identifying creation events (`ilrm:R18_created`) associated with a manuscript, which is further connected to a physical manifestation (`ilrm:R7i_is_materialized_in`). Each manifestation is linked to a creation event (`ilrm:R24_created`) that includes temporal information encoded as:

- a human-readable time-span label (`ecrm:P170i_time_is_defined_by`),
- a start date (`:has_start_date`),
- and an end date (`:has_end_date`).

To ensure historical precision, the query applies two `FILTER` clauses to both the start and end dates. These conditions allow only manuscripts with temporal boundaries that fall entirely within the extended range of 901–1100 CE to be included in the results. This inclusive approach accommodates chronological uncertainty and manuscripts that may straddle traditional century boundaries.

In addition to temporal data, the query retrieves bibliographic and archival metadata for each manuscript, including:

- the shelf mark or catalog identifier (`?shelfmark`),
- the number of folios (`?folios`),
- the name of the current holding library (`?libraryName`),
- and the name of the geographical location where the library resides (`?placeName`).

These details are collected by traversing relationships such as `ecrm:P50_has_current_keeper` (linking the manuscript to its repository), and `ecrm:P74_has_current_or_former_residence` (linking the library to a place). The final results are sorted chronologically and geographically by start and end date, place name, library name, and shelfmark, facilitating user navigation and comparative analysis.

By integrating temporal filtering with rich contextual metadata, this query provides an effective mechanism for the retrieval and study of 10th and 11th century manuscripts, supporting both historical research and digital cataloguing in the IMAGO knowledge base.

Table 9. Metadata about manuscripts produced within the 10th and 11th centuries.

Manuscript	Date
Gent, Universiteitsbibliotheek, 401 (152), 1r-21v	(sec. XI inizio)
Leiden, Universiteitsbibliotheek, Voss. lat. 4° 29, 1r-10r	(sec. XI inizio)
Montecassino, Archivio dell'Abbazia (Biblioteca Statale del Monumento Nazionale), 38, 128-140	(sec. XI inizio)
Alençon, Médiathèque de la Communauté Urbaine, 14, 1r-11v	(sec. XI)
Arras, Médiathèque de l'Abbaye Saint-Vaast, 832 (506), 1-117	(sec. XI)
Barcelona, Arxiu de la Corona d'Aragó, Ripoll 151, 1v-12r	(sec. XI)
Città del Vaticano, Biblioteca Apostolica Vaticana, Urb. lat. 440, 1r-8r	(sec. XI)
Città del Vaticano, Biblioteca Apostolica Vaticana, Vat. Lat. 13395, 83r-84v	(sec. XI)
Città del Vaticano, Biblioteca Apostolica Vaticana, Vat. lat. 4917, 146v-147v	(sec. XI)
Città del Vaticano, Biblioteca Apostolica Vaticana, Vat. lat. 4920, 48r-65v	(sec. XI)
Montecassino, Archivio dell'Abbazia (Biblioteca Statale del Monumento Nazionale), 152, 159-204	(sec. XI)
Namur, Bibliothèque du Séminaire Notre-Dame de Namur, 37, 118r-126r	(sec. XI)
Paris, Bibliothèque nationale de France, lat. 2321, 135v-143v	(sec. XI)
Paris, Bibliothèque nationale de France, lat. 5572, 112v-132v	(sec. XI)
Paris, Bibliothèque nationale de France, lat. 8518, 62-96	(sec. XI)
Paris, Bibliothèque nationale de France, n.a. lat. 1606, 20v-39v	(sec. XI)
Rouen, Bibliothèque Jacques Villon, A. 214 (469), 169r-173v	(sec. XI)
Saint-Omer, Bibliothèque d'Agglomération du Pays de Saint-Omer, 71, 101v-113v	(sec. XI)
Wien, Österreichische Nationalbibliothek, 563, 59v-112r	(sec. XI)
Wien, Österreichische Nationalbibliothek, lat. 580 , 10r-19v	(sec. XI)
Wolfenbüttel, Herzog August Bibliothek, Aug. 8° 56. 16 (3610), 1r-19r	(sec. XI)
Épinal, Bibliothèque Multimédia Intercommunale Epinal-Goldbey , 147 (67), 82v-95v	(sec. XI)
Montecassino, Archivio dell'Abbazia (Biblioteca Statale del Monumento Nazionale), 132, 1-530	(1022/1035)
Hamburg, Staats- und Universitätsbibliothek, philol. 122, 1r-4v	(sec. XI seconda metà)
London, British Library, Harley 2682, 180v-185r	(sec. XI seconda metà)
Venezia, Biblioteca Nazionale Marciana, lat. Z. 497 (1811), 183r-184r	(sec. XI seconda metà)
London, British Library, Royal 13 A I, 51v-78r	(sec. XI fine)

A Web Interface for Data Exploration

To address the second part of Research Question 2 (RQ2) - How can ontology-driven interfaces enable the exploration of a knowledge graph by users with different levels of expertise through simple interactions and user-friendly representations? - we implemented a web-based, ontology-driven interface for the IMAGO knowledge graph. The interface allows users to explore the knowledge graph by executing predefined queries without requiring knowledge (Pratelli et al. 2024b) of SPARQL or direct interaction with the SPARQL endpoint (Pratelli et al. 2024a).

The interface supports queries across six types of knowledge-extraction targets (see Figure 2), as described in the Knowledge Extraction and Evaluation via SPARQL Queries section, thereby accommodating users with varying levels of technical expertise and enabling intuitive, accessible exploration of the underlying semantic data.

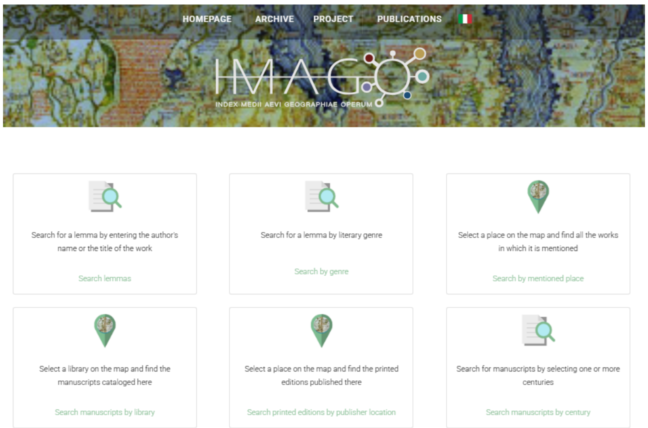

Figure 2. The web interface for the IMAGO KG exploration.

Query results are presented in user-friendly formats such as interactive tables and dynamic maps, making complex data more accessible and engaging for a broad audience (see Figure 3).

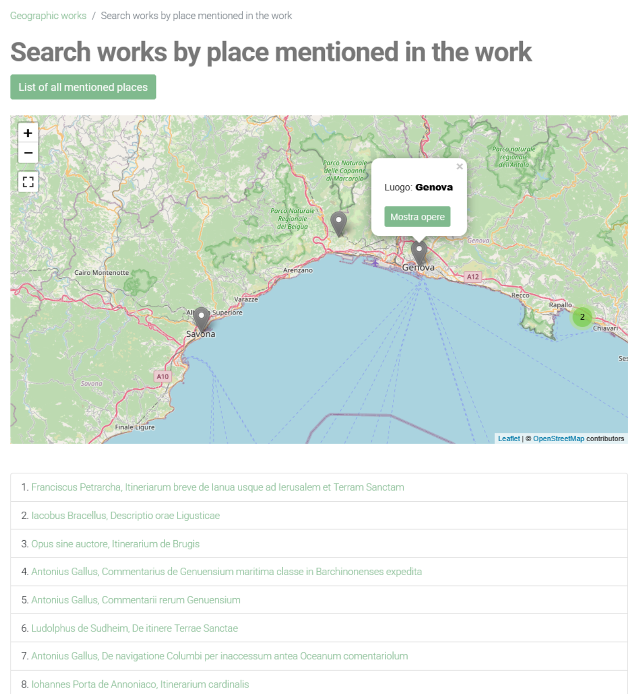

Figure 3. An example of the results of a query (Q3) showed on a map and as a table.

The interface is currently undergoing usability testing to assess and further improve the user experience. It is important to note that this user-based evaluation focuses specifically on the usability of the graphical interface, rather than on the effectiveness of queries or the research capabilities of the knowledge graph itself. The tasks proposed to users are designed to evaluate whether they can successfully complete predefined exploration activities - such as navigating the data, interpreting the results, and interacting with the system intuitively - rather than testing their ability to formulate domain-specific queries. Therefore, the evaluation emphasises ease of use, clarity of interaction, and overall user experience. In this sense, the GUI evaluation complements, but does not replace, the assessment of the KG's querying power, which is addressed separately through the definition and validation of domain-relevant queries, as described in the Knowledge Extraction and Evaluation via SPARQL Queries section.

Discussion

The IMAGO project presents significant achievements in the digital humanities field, offering the first comprehensive catalogue of medieval and Renaissance Latin geographical literature. By systematically collecting and organising these works, IMAGO provides scholars with an unprecedented resource for studying the geographical knowledge and perspectives of these historical periods. A key achievement of the project is the development of a standardised format for representing the collected dataset. By employing Semantic Web technologies and a Linked Open Data approach, and adhering to established and standard ontologies such as CIDOC CRM and LRMoo, the IMAGO ontology ensures semantic interoperability. This alignment facilitates data integration and sharing across various sources, enhancing the utility and reach of the IMAGO knowledge base. The structured nature of the IMAGO knowledge base enables automated data analysis, allowing for the extraction of relevant knowledge within the domain. This capability helps scholars uncover patterns and insights that might be difficult to discern through manual analysis. Furthermore, the predefined knowledge extraction queries are directly linked to higher-level scholarly tasks, such as investigating geographic references across multiple texts or tracing the transmission of works and their manuscripts over time. By integrating multiple dimensions of the data - author, place, genre, and manuscript location - these queries enable complex, semantically-informed searches that are difficult or time-consuming with traditional resources such as catalogues or critical editions. The relevance and utility of these queries were further validated through engagement with domain experts not involved in the design process, including internationally recognised scholars. Their feedback confirmed that the queries effectively support meaningful historical and philological investigations. Overall, the system accommodates both expert researchers and general users, providing accessible, ontology-driven tools for exploring the knowledge base through intuitive and domain-relevant interactions.

Despite its contributions, the IMAGO project faces certain limitations. While extensive, the dataset's current scope may not encompass all geographical works from the medieval and Renaissance periods. Additionally, the reliance on existing ontologies, while promoting interoperability, may constrain the representation of certain domain-specific nuances. Addressing these limitations will involve expanding the dataset and refining the ontology to capture a broader, more detailed range of information. Looking ahead, there are opportunities to expand the IMAGO knowledge base by incorporating additional knowledge and enhancing integration with related digital humanities projects, for example, the MMM project. Such efforts would further enrich the resource, providing a more comprehensive tool for scholars and facilitating interdisciplinary research. Continued development of the user interface, informed by ongoing usability studies, will also be crucial to ensuring that the platform remains accessible and responsive to its diverse user base.

As part of the discussion, it is also important to reflect on the evaluation of the proposed methodological pipeline, including the ontology-driven annotation tool. Rather than relying on large-scale automated benchmarks, the evaluation has been conducted along three complementary dimensions that are coherent with the expert-driven nature of the domain.

First, the pipeline ensures semantic consistency and data quality through the tight integration between the IMAGO ontology and the annotation tool. The use of controlled vocabularies, predefined entity lists, and semi-automatic IRI assignment constrains data entry and reduces ambiguity, while reasoning-based validation guarantees the logical consistency of the resulting knowledge graph.

Second, the effectiveness of the annotation tool has been assessed through its practical adoption within the IMAGO project. Domain experts were able to populate the knowledge graph with a substantial and complex corpus without requiring direct interaction with formal semantic technologies, confirming that the tool successfully mediates between domain expertise and knowledge representation.

Third, the evaluation of the pipeline is grounded in its ability to support meaningful knowledge extraction. The SPARQL queries defined in this work, derived from real scholarly questions, demonstrate that the resulting knowledge graph enables complex analyses that would be difficult to perform using traditional resources. The relevance of these queries has been further validated through feedback from domain experts.

Overall, while the proposed methodology does not aim at full automation, it demonstrates that an ontology-driven, expert-oriented pipeline can effectively support the creation, validation, and exploration of high-quality semantic data in the digital humanities domain.

Conclusion and Future Work

In this paper, we present the ontology and the knowledge graph developed within the IMAGO project, which aims to formally represent knowledge related to medieval and Renaissance Latin geographical works. To maximise interoperability, the IMAGO ontology was designed as an extension of the ISO standard CIDOC CRM.

To construct the knowledge graph, domain experts involved in the IMAGO project populated the ontology using a dedicated web-based annotation tool that we developed. This tool automatically assigns a unique IRI to each resource, including authors, manuscripts, literary genres, libraries, and places. To ensure alignment with existing semantic resources and promote interoperability, the tool retrieves IRIs from several freely available online sources, such as the Wikidata knowledge base, the MIRABILE digital archive, the Nuovo Soggettario thesaurus, and the Pleiades gazetteer. The resulting knowledge graph comprises 134,329 triples. It is published as an OWL 2 DL graph and stored in a Fuseki triplestore, which provides a SPARQL endpoint for querying the data.

However, beyond the release of the IMAGO knowledge graph, this work contributes a methodological pipeline for the semantic modelling, annotation, integration, and publication of data related to medieval and Renaissance geographical works, based on established Knowledge Representation and Semantic Web standards.

One of the main objectives of the project was the analysis of the collected data. To this end, through a series of plenary meetings involving the IMAGO community and project experts, we identified six key knowledge extraction targets. For each of these, we developed specific SPARQL queries to retrieve relevant information from the knowledge graph. These queries not only illustrate how the IMAGO knowledge graph can be used for data retrieval but also demonstrate its potential to support in-depth analysis and uncover new scholarly insights.

To make the data accessible and queryable by both domain experts and general users, a dedicated web application was developed. Through this interface, users can explore and query the knowledge base in a user-friendly and intuitive way. The application enables seamless exploration of the knowledge graph's content and fosters broader engagement with the project's results. The interface is currently undergoing usability testing to evaluate its effectiveness and further improve the overall user experience.

As future work, we plan to extend the evaluation of the IMAGO pipeline along several complementary directions. First, we aim to design and conduct more systematic experimental campaigns to assess the effectiveness of the methodological pipeline, including the annotation tool, in supporting domain experts during data acquisition. This will involve structured user studies with a broader set of scholars, focusing on both the quality of the produced annotations and the efficiency of the annotation process. In particular, the evaluation will consider quantitative usability metrics, including task completion rate, task completion time, error rates during annotation activities, and perceived ease of use, as well as comparative analyses of manual data entry and tool-assisted workflows.

Second, we plan to carry out a more formal evaluation of the knowledge graph and its querying capabilities, including the definition of additional competency questions. This will allow us to better quantify the benefits of the semantic approach in terms of information retrieval, integration, and analysis, as well as to assess query effectiveness and coverage with respect to domain research needs.

Another important direction concerns scalability. While the current knowledge graph already includes a substantial number of entities and triples, future work will investigate the behaviour of the pipeline when applied to larger and more heterogeneous corpora, as well as the performance of the triplestore and query execution under increasing data volumes. In parallel, scalability tests will assess system performance as additional manuscript catalogues and bibliographic resources are progressively ingested, with particular attention to query response times, reasoning tasks, and collaborative usage scenarios involving multiple domain experts.

In addition, we intend to further develop the web-based interface by incorporating feedback from ongoing usability studies and by introducing more advanced exploration features, such as flexible search mechanisms, faceted navigation, and support for more complex analytical queries.

Overall, these future developments will contribute to a more comprehensive assessment of the proposed approach and to the consolidation of the IMAGO framework as a robust and reusable infrastructure for digital humanities research.

Acknowledgements

This work was partially supported by ITSERR (Italian Strengthening of the ESFRI RI RESILIENCE), funded under the MUR National Recovery and Resilience Plan funded by the European Union-NextGenerationEU. The funders had no involvement in the study's design, data collection and analysis, publication decisions, or manuscript preparation.

References

Hubert Alisade, Diego Calvanese, Mario Klarer, Alessandro Mosca, Nonyelum Ndefo, Bernadette Rangger, and Aaron Tratter. Exploration of medieval manuscripts through keyword spotting in the MENS project. pages 67–74. CEUR-WS. URL https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-217345.

Konstantin Baierer, Evelyn Dröge, Kai Eckert, Doron Goldfarb, Julia Iwanowa, Christian Morbidoni, and Dominique Ritze. DM2E: A Linked Data source of Digitised Manuscripts for the Digital Humanities. *Semantic Web*, 8(5):733–745, April 2017. ISSN 22104968, 15700844. doi: 10.3233/SW-160234.

Valentina Bartalesi, Nicolò Pratelli, et al. The imago project: Towards a knowledge base of medieval and renaissance geographical works. In *CEUR WORKSHOP PROCEEDINGS*, 2021.

Valentina Bartalesi, Daniele Metilli, Nicolò Pratelli, and Paolo Pontari. Towards a knowledge base of medieval and renaissance geographical latin works: The imago ontology. *Digital Scholarship in the Humanities*, 37(1):34–50, 2022a.

Valentina Bartalesi, Nicolò Pratelli, and Emanuele Lenzi. Linking different scientific digital libraries in digital humanities: the imago case study. *International Journal on Digital Libraries*, 23(4):303–317, 2022b.

Toby Burrows. Manuscripts of Sir Thomas Phillipps in North American Institutions. *Manuscript Studies: A Journal of the Schoenberg Institute for Manuscript Studies*, 1(2):307–327, 2016. ISSN 2381-5329.

Toby Burrows. The History and Provenance of Manuscripts in the Collection of Sir Thomas Phillipps: New Approaches to Digital Representation. *Speculum*, 92(S1):S39–S64, October 2017. ISSN 0038-7134. doi: 10.1086/693438.

Toby Burrows. Connecting Medieval and Renaissance manuscript collections. *Open Library of Humanities*, 4(2), 2018. ISSN 2056-6700.

Toby Burrows, Doug Emery, Arthur Mitchell Fraas, Eero Hyvönen, Esko Ikkala, Mikko Koho, David Lewis, Andrew Morrison, Kevin Page, Lynn Ransom, Emma Cawlfield Thomson, Jouni Tuominen, Athanasios Velios, and Hanno Wijsman. A New Model for Manuscript Provenance Research: The Mapping Manuscript Migrations Project. *Manuscript Studies: A Journal of the Schoenberg Institute for Manuscript Studies*, 6(1):131–144, 2021. ISSN 2381-5329. doi: 10.1353/mns.2021.0004.

Paolo Chiesa. Per un riordino della tradizione manoscritta della" relatio" di odorico da pordenone. *Filologia mediolatina: rivista della fondazione Ezio Franceschini*, (6):311–352, 1999.

Domenico Defilippis. Da flavio biondo a leandro alberti. corografia e antiquaria tra quattro e cinquecento, atti del convegno di studi (foggia, 2 febbraio 2006), a cura di d. defilippis. 2009.

Mohamed Lamine Diakité, Béatrice Bouchou Markhoff, and Atigh Ould Boubacar. An ontology for Western Saharan manuscripts. In *2016 International Conference on Digital Economy (ICDEc)*, pages 60–67. doi: 10.1109/ICDEC.2016.7563146. URL https://ieeexplore.ieee.org/abstract/document/7563146.

Martin Doerr. The cidoc conceptual reference module: An ontological approach to semantic interoperability of metadata. *AI Mag.*, 24(3):75–92, September 2003. ISSN 0738-4602.

Martin Doerr, Stefan Gradmann, Steffen Hennicke, Antoine Isaac, Carlo Meghini, and Herbert Van de Sompel. The europeana data model (edm): object representations, context and semantics. Gothenburg; 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY, 2010.

Tom Elliott and Sean Gillies. Pleiades: The un-gis for ancient geography. *J. Geogr. Inf. Sci*, 22:1091–1108, 2008.

Charles B. Faulhaber and Óscar Perea Rodríguez. Philobiblon as a Digital Tool for Historians of Medieval Iberia. 2023. ISSN 1888-3931.

Mariano Fernández-López, Asunción Gómez-Pérez, and Natalia Juristo Juzgado. Methontology: from ontological art towards ontological engineering. 1997.

Sean Gilsdorf and Laura K. Morreale, editors. *Digital Medieval Studies—Experimentation and Innovation*. Arc Humanities Press, 2024. ISBN 978-1-80270-069-5. doi: 10.2307/jj.18108222.

Eero Hyvönen, Esko Ikkala, Jouni Tuominen, Mikko Koho, Toby Burrows, Lynn Ransom, and Hanno Wijsman. A Linked Open Data Service and Portal for Pre-modern Manuscript Research. *Digital Humanities in the Nordic and Baltic Countries Publications*, 2(1):220–229, May 2019. ISSN 2704-1441. doi: 10.5617/dhnbpub.11097.

Nancy Ide and Jean Véronis. *Text encoding initiative: Background and contexts*, volume 29. Springer Science & Business Media, 1995.

IFLA. International federation of library associations and institutions, 2023. https://www.ifla.org/.

Apache Jena. Apache jena fuseki. *The Apache Software Foundation*, 18, 2014.

Pietro Maria Liuzzo. Linked open data based on la syntaxe du codex for manuscripts in beta maṣāḥəft. *ISAW Papers*, 2021.

Anna Lucarelli and Giovanni Bergamin. The nuovo soggettario as a service for the linked data world. *JLIS: Italian Journal of*

Library, Archives and Information Science= Rivista italiana di biblioteconomia, archivistica e scienza dell'informazione: 4, 1, 2013, pages 213–226, 2013.

Béatrice Bouchou Markhoff, Mohamed Lamine Diakité, Sophie Caratini, Francesco Coreale, and Adel Ghamnia. Semantic web for biblimos. In *Extended Semantic Web Confenrece 2015, workshop Semantic Web For Scientific Heritage (SW4SH)*, 2015.

Alex Mayfield, Margaret Frei, Daryl Ireland, and Eugenio Menegon. The China Historical Christian Database: A Dataset Quantifying Christianity in China from 1550 to 1950. *Data*, 9(6):76, June 2024. ISSN 2306-5729. doi: 10.3390/data9060076.

Justin J Miller. Graph database applications and concepts with neo4j. In *Proceedings of the southern association for information systems conference, Atlanta, GA, USA*, volume 2324, pages 141–147, 2013.

Michel Mingam. Rameau, les catalogues, le web. 2015.

University of Erlangen-Nuremberg; Germanisches Nationalmuseum; Zoologisches Forschungsmuseum Koenig. Erlangen CRM / OWL, 2013. https://cidoc-crm.org/Resources/erlangen.

Openllet. Openllet: An Open Source OWL DL reasoner for Java, 2023. https://github.com/Galigator/openllet.

Lucia Pinelli. Mirabile : Condivisione del lavoro/condivisione dei saperi per lo studio del Medioevo. *Filologia medievale : comparatistica, critica del testo e attualità : atti del convegno (Viterbo, 26-28 settembre 2018)*, pages 195–201.

Paolo Pontari. Database ENSU. Edizione Nazionale dei testi della Storiografia umanistica. 2006.

Paolo Pontari, Elena Vagnoni, Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. Annotation template in english, 2026a. https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/docs/template_form_IT.doc.

Paolo Pontari, Elena Vagnoni, Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. Annotation template in italian, 2026b. https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/docs/template_form_IT.doc.

Paolo Pontari, Elena Vagnoni, Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. Annotation template in english, 2026c. https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/docs/filled-in_form_EN.docx.

Paolo Pontari, Elena Vagnoni, Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. Annotation template in english, 2026d. https://github.com/AIMH-DHgroup/IMAGO/blob/main/ontology/docs/filled-in_form_IT.doc.

Paolo Pontari et al. Nedum mille qui effluxerunt annorum gesta sciamus. l'italia di biondo e l'invenzione del medioevo. *SUPPLEMENTA HUMANISTICA LOVANIENSIA*, 39:151–175, 2016.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO SPARQL endpoint, 2024a. https://imagoarchive.it/sparql/.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO search interface, 2024b. https://imagoarchive.it/en/archivio.html.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. MMM Graph, 2026a. https://github.com/AIMH-DHgroup/IMAGO/blob/main/data/dumps/imago-mmm.ttl.gz.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO Main Graph, 2026b. https://github.com/AIMH-DHgroup/IMAGO/blob/main/data/dumps/imago-archive.ttl.gz.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. Imago knowledge graph before and after the validation process, 2026c. https://github.com/AIMH-DHgroup/IMAGO/tree/main/data/dumps.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO Knowledge Graph – Reasoning-based SPARQL Queries, 2026d. https://github.com/AIMH-DHgroup/IMAGO/blob/main/docs/Reasoning-based_SPARQL_Queries.md.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO Toponym Graph, 2026e. https://github.com/AIMH-DHgroup/IMAGO/blob/main/data/dumps/imago-toponyms.ttl.gz.

Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. IMAGO Triplifier, 2026f. https://github.com/AIMH-DHgroup/IMAGO/tree/main/tools/triplifier.

Pat Riva and Maja Žumer. Frbroo, the ifla library reference model, and now lrmoo: a circle of development. In *IFLA WLIC 2018 Conference, Transform Libraries, Transform Societies*, 2017.

Pat RIva, Maja Žumer, and Trond Aalberg. LRMoo, a high-level model in an object-oriented framework. 2022.

Luigi Russo. ALIM, Archivio della latinità italiana del Medioevo. *Reti Medievali Rivista*, 6(1):5–5, June 2005. ISSN 1593-2214. doi: 10.6092/1593-2214/181.

Denny Vrandečić and Markus Krötzsch. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, September 2014. ISSN 0001-0782. doi: 10.1145/2629489. URL https://doi.org/10.1145/2629489.

Appendix A

In the following, the code of the six predefined queries is reported (the code is also available on GitHub: https://github.com/AIMH-DHgroup/IMAGO/blob/main/docs/Six-SPARQL-Queries.txt).

To facilitate testing and ensure the reproducibility of the results, we have also created six dedicated pages on the IMAGO platform, one for each query, with the query preloaded. These pages allow users to run the queries against our SPARQL endpoint instantly, with a single click. The query pages are listed below:

- Q1: https://imagoarchive.it/sparql/example-queries/q1.html
- Q2: https://imagoarchive.it/sparql/example-queries/q2.html
- Q3: https://imagoarchive.it/sparql/example-queries/q3.html
- Q4: https://imagoarchive.it/sparql/example-queries/q4.html
- Q5: https://imagoarchive.it/sparql/example-queries/q5.html
- Q6: https://imagoarchive.it/sparql/example-queries/q6.html

Q1. The SPARQL query retrieves author–work pairs (i.e. lemma) using name and title input.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT ?exp_cre ?title ?authorName ?alias
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  ?exp_cre a ilrm:F28_Expression_Creation ;
           ilrm:R17_created ?work ;
           ecrm:P14_carried_out_by ?author .

  ?author a :Author ;
          ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?authorName ;
          ecrm:P1_is_identified_by/:has_alias ?alias .

  ?work a ilrm:F2_Expression ;
        ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?title .

  FILTER (?title = "Descriptio insule Crete")
  FILTER (?authorName = "Christophorus Bondelmontius")
}
ORDER BY ?authorName ?title
```

Q2. The query retrieves the works classified under the Cartography literary genre. For each work, the title and author name are extracted. The genre is identified by the IRI `https://imagoarchive.it/thes/tid/10003`.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT ?exp_cre ?title ?authorName
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  BIND(<https://imagoarchive.it/thes/tid/10003> AS ?genre)

  ?exp_cre a ilrm:F28_Expression_Creation ;
           ilrm:R17_created ?work ;
```

```
           ecrm:P14_carried_out_by ?author .

  ?author a :Author ;
          ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?authorName .

  ?work a ilrm:F2_Expression ;
        ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?title ;
        :has_genre ?genre .
}
ORDER BY ?authorName
```

Q3. The query retrieves all works in which the city of Genoa is mentioned. The query also extracts the title of the work and the author's name, and links the toponym to its geographic representation.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT DISTINCT ?exp_cre ?authorName ?title
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  BIND(<https://imagoarchive.it/ontology/resources/toponym/genova> AS ?toponym)

  ?exp_cre a ilrm:F28_Expression_Creation ;
           ilrm:R17_created ?work ;
           ecrm:P14_carried_out_by ?author .

  ?author a :Author ;
          ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?authorName .

  ?work a ilrm:F2_Expression ;
        ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?title ;
        ecrm:P106_is_composed_of ?toponym .

  ?place :is_identified_by_toponym ?toponym ;
         ecrm:P168_place_is_defined_by/ecrm:P190_has_symbolic_content ?coord .

  ?toponym ecrm:P190_has_symbolic_content ?placeName .
}
```

Q4. The SPARQL query retrieves manuscripts currently held by the Biblioteca Nazionale di Firenze. The query returns information on each manuscript's shelfmark, number of folios, holding institution, and associated geographic location.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT ?manuscript ?placeName ?libraryName ?shelfmark ?folios
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  BIND(<http://www.wikidata.org/entity/Q460907> AS ?library)

  ?exp_cre a ilrm:F28_Expression_Creation ;
```

```
          ilrm:R17_created ?work ;
          ecrm:P14_carried_out_by ?author .

  ?author a :Author ;
         ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?authorName .

  ?work a ilrm:F2_Expression ;
       ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?title .

  ?exp_cre ilrm:R18_created ?manuscript .

  ?manuscript
    ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?shelfmark ;
    ecrm:P50_has_current_keeper ?library ;
    ecrm:P46_is_composed_of/ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content
    ?folios .

  ?manifestation ilrm:R7i_is_materialized_in ?manuscript .

  ?library
    ecrm:P74_has_current_or_former_residence ?libraryPlace ;
    ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?libraryName .

  ?libraryPlace
    :is_identified_by_toponym ?toponym ;
    ecrm:P168_place_is_defined_by/ecrm:P190_has_symbolic_content ?coord .

  ?toponym ecrm:P190_has_symbolic_content ?placeName .
}
ORDER BY ?placeName ?libraryName ?shelfmark
```

Q5. The query retrieves all printed editions of manuscripts published in Bologna. The query includes information about the edition's title, author, place of publication, and - where available - publisher, date, and curator.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT ?print_author ?print_edition ?print_title ?l_datazione ?placeName ?publisher
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  BIND(<http://www.wikidata.org/entity/Q1891> AS ?toponym)

  ?exp_cre a ilrm:F28_Expression_Creation ;
          ilrm:R17_created ?work ;
          ecrm:P14_carried_out_by ?author .

  ?author a :Author ;
         ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?authorName .

  ?work a ilrm:F2_Expression ;
       ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?title .

  ?printEditionCreation ilrm:R24_created ?print_edition .

  ?print_edition
    ilrm:R4_embodies ?work ;
```

```
    :is_composed_of_author/ecrm:P190_has_symbolic_content ?print_author ;
    ecrm:P102_has_title/ecrm:P190_has_symbolic_content ?print_title .

  ?print_creation ilrm:R24_created ?print_edition ;
                  ecrm:P7_took_place_at ?toponym .

  ?toponym :is_identified_by_toponym/ecrm:P190_has_symbolic_content ?placeName .

  OPTIONAL {
    ?printEditionCreation
      ecrm:P4_has_time-span/ecrm:P170i_time_is_defined_by ?l_datazione .
  }

  OPTIONAL {
    ?print_creation :has_curator/ecrm:P1_is_identified_by
    /ecrm:P190_has_symbolic_content ?curator .
  }

  OPTIONAL {
    ?print_creation :has_publisher/ecrm:P1_is_identified_by
    /ecrm:P190_has_symbolic_content ?publisher .
  }
}
```

Q6. The query retrieves the data associated (i.e., library, place in which the library is located, and shelf mark) with the manuscripts dated to the 10^{th} and 11^{th} century.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ecrm: <http://erlangen-crm.org/211015/>
PREFIX ilrm: <http://imagoarchive.it/ilrmoo/>
PREFIX : <https://imagoarchive.it/ontology/>

SELECT DISTINCT ?manuscript ?placeName ?libraryName ?shelfmark
                ?folios ?date_manuscript ?start_date_manuscript
                ?end_date_manuscript
FROM <https://imagoarchive.it/fuseki/imago/archive>
WHERE {
  ?exp_cre ilrm:R18_created ?manuscript .
  ?manuscript
    ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?shelfmark ;
    ecrm:P50_has_current_keeper ?library ;
    ecrm:P46_is_composed_of/
        ecrm:P1_is_identified_by/
            ecrm:P190_has_symbolic_content ?folios .

  ?manifestation ilrm:R7i_is_materialized_in ?manuscript .
  ?manifestation_creation
    ilrm:R24_created  ?manifestation ;
    ecrm:P4_has_time-span/ecrm:P170i_time_is_defined_by ?date_manuscript ;
    :has_start_date ?start_date_manuscript ;
    :has_end_date ?end_date_manuscript .

  ?library
    ecrm:P74_has_current_or_former_residence ?libraryPlace ;
    ecrm:P1_is_identified_by/ecrm:P190_has_symbolic_content ?libraryName .

  ?libraryPlace :is_identified_by_toponym ?toponym .
  ?toponym ecrm:P190_has_symbolic_content ?placeName .
```

```
  FILTER((
    "901-01-01T00:00:00Z"^^xsd:dateTime <= ?start_date_manuscript &&
    ?start_date_manuscript <= "1000-01-01T00:00:00Z"^^xsd:dateTime
  ) || (
    "1001-01-01T00:00:00Z"^^xsd:dateTime <= ?start_date_manuscript &&
    ?start_date_manuscript <= "1100-01-01T00:00:00Z"^^xsd:dateTime
  ))

  FILTER((
    "901-01-01T00:00:00Z"^^xsd:dateTime <= ?end_date_manuscript &&
    ?end_date_manuscript <= "1000-01-01T00:00:00Z"^^xsd:dateTime
  ) || (
    "1001-01-01T00:00:00Z"^^xsd:dateTime <= ?end_date_manuscript &&
    ?end_date_manuscript <= "1100-01-01T00:00:00Z"^^xsd:dateTime
  ))
}
ORDER BY ?start_date_manuscript ?end_date_manuscript ?placeName ?libraryName ?shelfmark
```