

# Using Semantic Web to Create and Explore an Index of Toponyms Cited in Medieval Geographical Works

VALENTINA BARTALESI\*, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), Italy

NICOLO' PRATELLI\*, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), Italy and Department of Information Engineering, University of Pisa, Italy

EMANUELE LENZI\*, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), Italy

PAOLO PONTARI, Department of Philology, Literature and Linguistics - University of Pisa, Italy

Western thought in European history was mainly affected by the image of the world created during the Middle Ages and Renaissance. The most popular reason to travel during the Middle Ages was taking a pilgrimage. Jerusalem, Rome, and Santiago de Compostela were the most popular destinations. It is not surprising that a lot of works written by travellers as guides for pilgrims exist. By the beginning of the Renaissance, a more precise image of the world was defined thanks to the discovery of ancient geographical models, especially the work of Ptolemy. The three years (2020-2023) Italian National research project IMAGO - Index Medii Aevi Geographiae Operum - aims to provide a systematic overview of the medieval and renaissance Latin geographical literature using the Semantic Web technologies and the LOD paradigm. Indeed, until now, this literature has not been studied using digital methods. In particular, this paper presents how we formally represented the knowledge about the toponyms, or place names, in the IMAGO ontology. To maximise the interoperability, we developed the IMAGO ontology as an extension of two reference vocabularies: the CIDOC CRM and its extension FRBRoo, including its in-progress reformulation, LRMoo. Furthermore, we used Wikidata as reference knowledge base. As case study, we chose to represent the knowledge related to the toponyms cited by the Italian poet Dante Alighieri in his Latin works. We carried out a first experiment for visualising the knowledge about these toponyms on a map and in the form of tables and CSV files.

Additional Key Words and Phrases: Semantic Web, Linked Open Data, ontology, toponyms, CIDOC CRM, Wikidata, Dante Alighieri

## ACM Reference Format:

Valentina Bartalesi, Nicolo' Pratelli, Emanuele Lenzi, and Paolo Pontari. 2022. Using Semantic Web to Create and Explore an Index of Toponyms Cited in Medieval Geographical Works. *ACM J. Comput. Cult. Herit.* 37, 4, Article 111 (August 2022), 18 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

\*Both authors contributed equally to this research.

Authors' addresses: Valentina Bartalesi, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), via G. Moruzzi 1, Pisa, , Italy, 56124, [valentina.bartalesi@isti.cnr.it](mailto:valentina.bartalesi@isti.cnr.it); Nicolo' Pratelli, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), via G. Moruzzi 1, Pisa, , Italy, 56124 and Department of Information Engineering, University of Pisa, Via G. Caruso 16, Pisa, Italy, [nicolo.pratelli@isti.cnr.it](mailto:nicolo.pratelli@isti.cnr.it); Emanuele Lenzi, Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), via G. Moruzzi 1, Pisa, , Italy, 56124, [emanuele.lenzi@isti.cnr.it](mailto:emanuele.lenzi@isti.cnr.it); Paolo Pontari, Department of Philology, Literature and Linguistics - University of Pisa, piazza Evangelista Torricelli, 2, Pisa, Italy, [paolo.pontari@unipi.it](mailto:paolo.pontari@unipi.it).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2022 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104



Fig. 1. Matthew Paris's Map of the Route to Jerusalem. St. Albans, c.1250. London, British Library, Royal MS, 14 C vii, f. 5.

## 1 INTRODUCTION

The world's image created during the Middle Ages and Renaissance was crucial to the development of Western thought in European history. Especially in the Middle Ages, most people did not travel, and their knowledge and experience of the world were restricted to their immediate location. The most popular reason to travel was taking a pilgrimage. Jerusalem, Rome, and Santiago de Compostela were the most popular destinations. These trips would take months and require elaborate planning [17]. It is not surprising that there are works written by travellers as guides for pilgrims [25]. These guides report specific descriptions of lands, cities, places, monuments and buildings, as well as advice about what to pack, what to wear, and where and how to sleep [1]. An example of these guides is Matthew Paris's itinerary maps from London to Palestine. Around 1250, Mathew Paris, a monk, chronicled events from the creation of the world to the mid 13th century in his work *Chronica Major*. At the beginning of a manuscript of this work, we find a map showing the pilgrimage route from London to Jerusalem (Figure 1). This medieval road map is spread over seven pages and takes the viewer on a journey from London through significant stops on the pilgrimage [10].

By the end of the Middle Ages and the beginning of the Renaissance, a more precise image of the world was defined thanks to the discovery of ancient geographical models, especially the work of Ptolemy. Ptolemy was a Greco-Roman geographer, author of *Geographia*, a work written about AD 150, in which all of the knowledge about the world's geography in the Roman Empire of the 2nd century was reported and described. In the 15th century, his work was rediscovered and translated into Latin from the original Greek. The influence of Ptolemy on the medieval cartography is due to his use of applying geographic coordinates to 8,000 geographic locations, as well as the use of mathematical calculations to depict geographical places. The genre of geographical description had a further and fundamental turning point during the period between the mid-15th to the mid-16th century, when a combination of circumstances encouraged people to seek new routes and to start the *Age of Discovery*. Indeed, toward the end of the 14th century, the empire of the Mongols had been broken up, and Western merchants could no longer be assured of safe conduct along the land routes.

105 Furthermore, the Ottoman Empire and the Republic of Venice controlled commercial access to the Mediterranean and  
106 the ancient sea routes from the East. Thus, nations, especially on the Atlantic shores of Europe, started to seek overseas  
107 trade and explorations. These travels and discoveries allowed the reassessment of the physical space and the description  
108 and representation of the new routes and lands, and the New World gave the basis of modern geography [11].  
109

110 The three years (2020-2023) Italian National research project IMAGO - Index Medii Aevi Geographiae Operum - aims  
111 to provide a systematic overview of the medieval and renaissance Latin geographical literature using the Semantic  
112 Web technologies [18]. Indeed, until now, this literature has not been studied using digital methods. The final aim of  
113 the project is the creation of a Web application allowing scholars to freely access and visualise the data collected in  
114 the IMAGO knowledge base and to publish this knowledge as Linked Open Data (LOD) [6] and Findable, Accessible,  
115 Interoperable, and Reusable (FAIR) data [39]. In particular, the project wants to provide: (i) a collection of the manuscript  
116 tradition and printed editions for each work; (ii) a classification of authors, genres and contents; (iii) a collection of  
117 critical editions of some more representative works; (iv) a medieval Latin toponym index. As the first step to developing  
118 tools to support scholars in creating, evolving and consulting a knowledge base (KB) of the geographical works, we  
119 created an OWL 2 DL ontology [23] that formally represents this knowledge. OWL 2 DL guarantees the decidability  
120 of the language, and it has a sufficient representational power for the project aims. Following the re-use logic and to  
121 maximise interoperability, we developed our ontology as an extension of two reference ontologies, that is, the CIDOC  
122 CRM vocabulary [13] and its extension FRBRoo [14], including its in-progress reformulation, LRMoo [32]. We have  
123 also designed and developed a Web tool that allows scholars to populate the ontology. Then, we carried out a first  
124 experiment for visualising the knowledge about Dante Alighieri's Latin toponyms collected in our KB on a map and  
125 in the form of tables and CSV files. To guarantee the quality of the collected dataset, all data were produced by the  
126 authoritative scholars involved in the project. After the creation of the KB, two or more scholars manually checked all  
127 data for consistency to guarantee the absence of systematic errors.  
128

129 This paper presents the process we followed to formally represent the knowledge about the toponyms in the IMAGO  
130 ontology (Section 3). To contextualise our work, in Section 2, we report an overview of the current datasets of  
131 geographical places that use Semantic Web technologies and the LOD paradigm to facilitate the interlinking of different  
132 sources. Section 4 reports a description of the semi-automatic tool we designed and developed to allow scholars to  
133 populate the ontology. In Section 5, we describe the visualisation components to access and explore the knowledge  
134 about the toponyms. As a case study, we chose to represent and visualise the knowledge related to the toponyms cited  
135 by the Italian poet Dante Alighieri in his Latin works. This choice is motivated by the fact that the ontology population  
136 phase is far from the end, and in this first period, the scholars involved in our project focused their research on Dante  
137 Alighieri's works. Finally, in Section 6, some concluding remarks are given.  
138

## 144 2 RELATED WORKS

146 Medieval and renaissance Latin geographical works report many place names or toponyms and natural descriptions.  
147 Having toponyms in a standardised digital form allows scholars to compare their toponym datasets with other datasets,  
148 enrich their datasets with additional knowledge, and produce aggregated views on the data, e.g. aggregation of current  
149 and historical information about a place. Semantic Web technologies and the LOD paradigm facilitate the sharing  
150 and the comparison of the data. A dataset is linked and open when it is structured and published according to the  
151 LOD principles [7], so that it can be both interlinked and made openly accessible and shareable on the Semantic Web.  
152 Indeed, the goal of LOD is to allow data from different resources to be interconnected and queried. To make this data  
153 understood and automatically processed by computers, it must be expressed in a standard format. In the Semantic Web,  
154  
155

each entity (for example, a museum object, a place, or a person) must be represented by a persistent identifier (known as an Internationalised Resource Identifier (IRI)<sup>1</sup>). Furthermore, RDF/S (Resource Description Framework/Schema) [27] and OWL (Web Ontology Language) [24] are formal languages for describing entities, as well as the relationships between them as simple properties and values (known as *triples*). The Semantic Web and LOD approaches are currently used to represent knowledge about toponyms. In the following, we report some well-known datasets of geographical places that use Semantic Web technologies and the LOD paradigm to facilitate the interlinking of different sources.

The first example is the Getty Thesaurus of Geographic Names<sup>2</sup> (TGN). It is a structured vocabulary containing over four million toponyms and associated information about places. Places include administrative entities (e.g., cities, nations) and physical features (e.g., mountains, rivers). Other information related to history, population, culture, art and architecture is included. While many records in TGN include geographic coordinates, these coordinates are approximate and are intended for reference only. Indeed, they typically represent a single point, corresponding to a point in or near the centre of the inhabited place, political entity, or physical feature. TGN is constructed to allow publishing its data as LOD, i.e. JSON, RDF, N3/Turtle, N-Triples.

Another widely adopted open dataset for geographical entities is the Geonames Geographical Database<sup>3</sup>. It contains over 27 million geographical names. In addition to the names of places in various languages, Geonames includes latitude, longitude, elevation, population, administrative subdivision and postal codes. The geographical coordinates, stored with all the names, enable combining the municipality with the place names used for an inhabited place that is not an administrative unit. As a result of this spatial combination, the web pages provide lists of all known settlements in a municipality. Each GeoNames feature is represented as a web resource identified by an IRI. This IRI provides access either to the HTML wiki page or to an RDF description of the feature, using elements of the GeoNames ontology<sup>4</sup>. This ontology describes the GeoNames features properties using the Web Ontology Language, the feature classes and codes being described in the SKOS language.

LinkedGeoData [12] is another effort to add a spatial dimension to the Web of Data. LinkedGeoData uses the information collected by the OpenStreetMap (OSM) project<sup>5</sup> and makes it available as an RDF knowledge graph according to the Linked Data principles. LinkedGeoData is based on an ontology that is derived from the OSM tags [35]. The ontology includes around 1,200 classes, 250 data properties, and 80 object properties.

Another example of a toponyms dataset is Wikidata [38], an open collaborative general-purpose knowledge base created by the Wikimedia Foundation. The knowledge base accepts editing by any user, following the model of Wikipedia. Wikidata currently contains more than 17 million items and thousands of descriptions of geographical entities. For each geographical entity, Wikidata reports the latitude and longitude of single points. Wikidata has reached full compatibility with Semantic Web technologies [15], and it provides a SPARQL endpoint to query the knowledge base. Since October 2014, a GeoNames ID external identifier has been reported for many geographical Wikidata entities.

Pleiades<sup>6</sup> gazetteer provides historical geographical information about the ancient world in digital form. At present, Pleiades has extensive coverage for the Greek and Roman world and is expanding into Ancient Near Eastern, Byzantine, Celtic, and Early Medieval geography. Pleiades is compliant with the Semantic Web technologies and uses terms from

<sup>1</sup><https://www.w3.org/International/articles/idn-and-iri/>

<sup>2</sup><http://www.getty.edu/research/tools/vocabularies/tgn/index.html>

<sup>3</sup><http://www.geonames.org/>

<sup>4</sup><http://www.geonames.org/ontology/documentation.html>

<sup>5</sup><https://www.openstreetmap.org>

<sup>6</sup><https://pleiades.stoa.org/home>

209 different vocabularies to represent the collected knowledge, e.g. Dublin Core, SKOS, FOAF<sup>7</sup>, Citation Ontology<sup>8</sup>. The  
210 Pleiades dataset may be downloaded in RDF format.

211 The project Mapping Manuscript Migrations (MMM) makes available a dataset of places relating to the history and  
212 provenance of Western European medieval and early modern manuscripts [9]. The MMM links disparate datasets  
213 from Europe and North America to provide a view of the provenance of medieval and renaissance manuscripts. The  
214 MMM data are available as LOD for reuse under a CC-BY-NC 4.0 licence. Another project that makes available a  
215 geographic dataset as LOD is HiGeoMes. This project aims to integrate information on places mentioned in Babylonian  
216 and Assyrian texts with archaeological sites [21]. The linking of documented settlements with place names from written  
217 sources is intended to provide a better understanding of political, social and environmental developments in Upper  
218 Mesopotamia in the 2nd millennium BC. Finally, we want to mention an ongoing project LINKEDOPENNEPAL, which  
219 aims at creating an ontology for representing the Nepalese toponyms [37] attested in the documents collected in a  
220 text corpus that is part of the Documenta Nepalica<sup>9</sup>. This corpus comprises texts and documents on the history of  
221 religion and law of pre-modern Nepal. The ontology extends CIDOC CRM and increases interoperability by aligning  
222 the ontology individuals to the respective entries of GeoNames. Also, a mapping of the individuals to DBpedia entities  
223 was established.

224 As reported in detail in the following Section, to represent the knowledge collected in the IMAGO project, we defined an  
225 ontology as an extension of the standard vocabulary CIDOC-CRM, and its extension FRBRoo, including its in-progress  
226 reformulation, LRMoo. Indeed, the CRM provides terms to represent places, toponyms and geographic coordinates. In  
227 this representational context, we used Wikidata as reference knowledge base to represent the geographical entities we  
228 found in the works collected in the IMAGO project. We decided to use the IRIs of Wikidata since it is a general-purpose  
229 KB; thus, we could also find IRIs for other entities (e.g. authors, works, libraries) from the same source.

### 236 3 THE IMAGO ONTOLOGY

#### 237 3.1 General Description of the Ontology

238 Despite in other research projects semantic technologies have been used to represent ancient manuscript corpora  
239 [5, 9, 16, 20], no scientific research that applies a Semantic Web approach has been conducted in the medieval and  
240 renaissance Latin geographical literature. Indeed, the main novelty introduced by our research is the use of Semantic  
241 Web technologies to formally represent this scientific domain. The knowledge related to these ancient works is currently  
242 dispersed on paper books. This makes a systematic overview of this geographical literature impossible, preventing a  
243 well-ordered perception of how it was gradually set up in time. The IMAGO project aims at making this information  
244 easily available in digital form for scholars, students and general users.

245 As the first step to develop tools to support scholars in creating, evolving and consulting a KB of the medieval and  
246 renaissance geographical works, we created an ontology that formally represents this knowledge. The ontology is  
247 extensively described in [2] and [3]. The IMAGO ontology is derived from a conceptualisation of the knowledge  
248 domain, which identifies the categories and the relations that have to be formally represented. This conceptualisation  
249 is the result of the collaboration between the CNR and the experts in Latin geographical literature who are involved  
250 in the project, i.e. scholars from the University of Pisa and the University of Salento<sup>10</sup>. The scholars identified the

256 <sup>7</sup><http://xmlns.com/foaf/0.1/>

257 <sup>8</sup><http://purl.org/spar/cito/>

258 <sup>9</sup><https://nepalica.hadw-bw.de/nepal/>

259 <sup>10</sup><https://imagoarchive.it/gruppo.html>

261 categories and relationships of the knowledge domain they were interested in formally representing. In particular, the  
 262 conceptualisation is based on authoritative studies of the medieval and renaissance Latin geographical works, i.e. [28],  
 263 [29], [25], [22], [26], [11], [8]. In our conceptualisation, we identified some main categories that represent the domain  
 264 of the geographical works. The first categories are the author and title of a work. For each work, the literary genre is  
 265 specified along with the toponyms that represent the places that are described or reported in the work. Furthermore,  
 266 for each work, several metadata about the related manuscripts and printed editions are defined. In particular, for each  
 267 manuscript, the following categories are identified:  
 268  
 269

- 270 • the name of the author and the title of the work in the forms that appear in the manuscript;
- 271 • the library in which the manuscript is collected;
- 272 • the location of the library;
- 273 • the signature and the folios of the manuscript;
- 274 • the *incipit* and *explicit* of the dedication/proem, if they exist;
- 275 • the *incipit* and *explicit* of the text;
- 276 • the date of the creation of the manuscript;
- 277 • the secondary sources;
- 278 • the iconographic apparatus;
- 279 • the link to the description or reproduction of the manuscript;
- 280 • the scholars' notes.

281  
 282  
 283  
 284 On the other hand, for each printed edition, the following categories are defined:

- 285 • the author's name and the title as reported in the edition;
- 286 • the curator's name;
- 287 • the place and the date of publication;
- 288 • the publisher;
- 289 • the format of the edition;
- 290 • the number of pages;
- 291 • the information about the images reported in the edition;
- 292 • some general notes that the scholars intend to add to the edition;
- 293 • the author of the introduction;
- 294 • the text of the introduction;
- 295 • the text of the dedications;
- 296 • information about whether the edition is a first edition or a reprint;
- 297 • the primary and secondary sources of the edition;
- 298 • and the ecdotic typology.

303  
 304 Once the conceptualisation was defined, we formalised it using classes and properties from two existing ontologies we  
 305 chose as reference vocabularies, i.e. the CIDOC CRM and its extension FRBRoo, including its in-progress reformulation  
 306 LRMoo. We adopted a lot of terms from these ontologies to maximise the interoperability of our representation. Where  
 307 necessary, we added our own classes and properties to represent the terms that we did not find in the reference  
 308 vocabularies. The resulting ontology is an extension of the CRM, and it is expressed in OWL 2 DL [23]. In Table 1, we  
 309 report the classes that represent our main concepts and in Table 2, the properties to express the semantic relations  
 310 among concepts are listed. As a notational convention, the CIDOC CRM uses the letters "E" and "P" to indicate classes  
 311

and properties, respectively, whereas FRBRoo (and its revisions LRMoo) uses the letters "F" and "R" to indicate classes and properties, respectively.

Table 1. Classes used to represent the IMAGO main concepts

Concept	Class
Author	subclass of E39 Actor
Work	equivalent to F2 Expression
Work creation	equivalent to F28 Expression Creation
Genre	subclass of E55 Type
Toponym	subclass of E41 Appellation
Manuscript	subclass of F5 Item
Printed Edition	subclass of F3 Manifestation
Library	subclass of F11 Corporate Body
Place	equivalent to E53 Place
Geographic Coordinate	equivalent to E94 Space Primitive
Signature	equivalent to E42 Identifier
Folios	subclass to E19 Physical Object
Date	equivalent to E52 Time-Span
Curator/Publisher	subclass of E39 Actor

Table 2. Properties used to represent the relations among the IMAGO main concepts

Relation (R) between concepts	Property
R(Work creation event,Author)	equivalent to P14 is carried out by
R(Work creation event,Work)	equivalent to R17 created
R(Manuscript,Title)	equivalent to P102 has title
R(Printed Edition,Title)	equivalent to P102 has title
R(Manuscript,Library)	equivalent to P50 has current keeper
R(Place,Geographic coordinates)	equivalent to P168 place is defined by
R(Manuscript,Signature)	equivalent to P1 is identified by
R(Manuscript,Folios)	equivalent to P46 is composed of
R(Manuscript,Date)	equivalent to P4 has time span
R(Printed edition,Date)	equivalent to P4 has time span
R(Printed edition,Curator)	subproperty of P14 carried out by
R(Printed edition,Publisher)	subproperty of P14 carried out by
R(Printed edition,Format)	equivalent to R69 specifies physical form
R(Printed edition,Page)	equivalent to P106 is composed of

Figure 2 shows the main classes and properties of the IMAGO ontology.

Following the Semantic Web approach, all resources stored in the IMAGO KB have a unique identifier, i.e. the Internationalized Resource Identifier (IRI). These IRIs are automatically retrieved from different sources freely available on the Web. In particular, we used the IRIs stored in the following datasets: (i) Wikidata knowledge base [38]; (ii) MIRABILE digital archive<sup>11</sup>; (iii) Mapping Manuscript Migrations knowledge base[19]; and (vi) Plaiades gazeteer<sup>12</sup>. When the IRIs were not available in the previous datasets, we automatically created and assigned custom IRIs to the resources.

We were able to retrieve 96% of author IRIs and 98% of work IRIs from Wikidata and Mirabile. 95% of place IRIs and 41%

<sup>11</sup>[www.mirabileweb.it](http://www.mirabileweb.it)

<sup>12</sup><https://pleiades.stoa.org/>

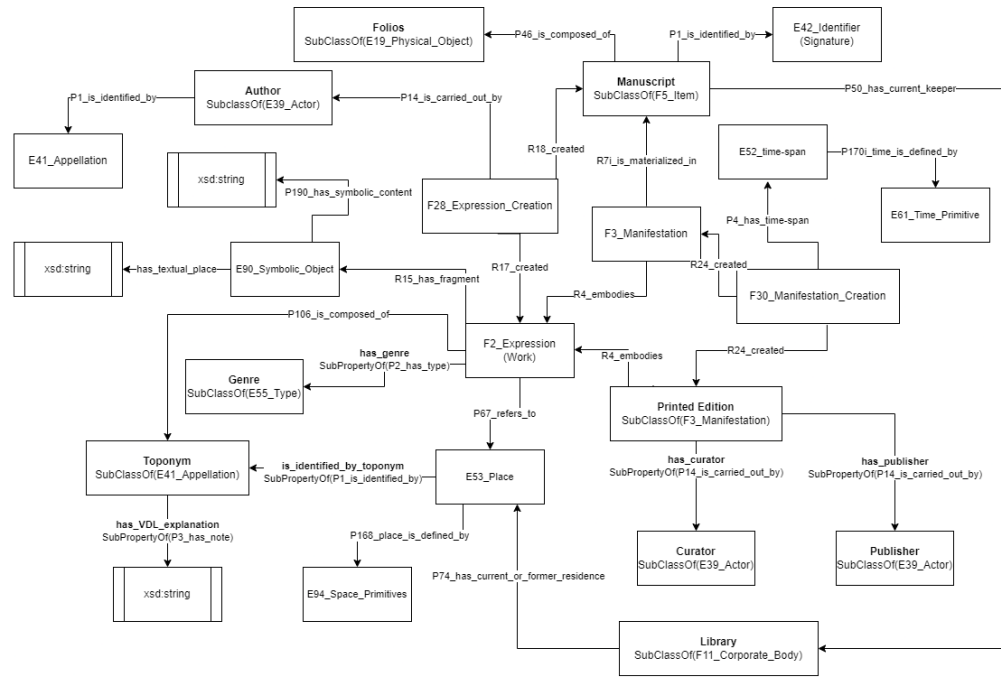


Fig. 2. The main classes and properties of the IMAGO ontology.

of library IRIs from Wikidata. 50% of ancient place IRIs from Pleiades. Finally, we were able to map 20% of the works collected in the IMAGO project with those collected in the MMM KB.

Linking our KB to these different datasets allows significantly enriching the knowledge collected in the IMAGO project. Indeed, each resource has an associated description reporting new knowledge in comparison with the one collected by the scholars participating in the IMAGO project.

### 3.2 Representing Toponyms in the IMAGO Ontology

One of the aims of our ontology is the representation of the knowledge related to the toponyms that are reported in the medieval works we collected. In particular, the ontology is able to represent the following pieces of knowledge:

- the place identified by the toponym;
- the geographic coordinates associated with the place identified by the toponym;
- the work/s in which the toponym appears;
- the textual fragment in which the toponym appears;
- the textual place of the toponym;
- the author of the work in which the toponym is reported.

In the following, we report a running example to better clarify the meaning of the pieces of knowledge reported above. To test our model, we started to populate the ontology with the toponyms, and the related data, cited in the Latin works of the Italian poet Dante Alighieri<sup>13</sup>. The works we took into account are *De Vulgari Eloquentia*, *Monarchia*,

<sup>13</sup>[https://en.wikipedia.org/wiki/Dante\\_Alighieri](https://en.wikipedia.org/wiki/Dante_Alighieri)



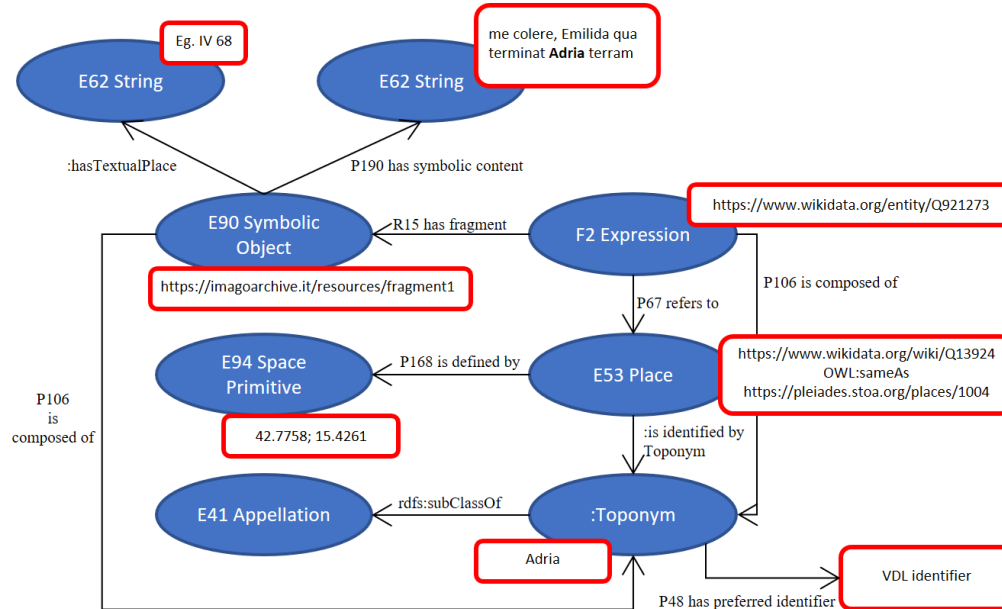


Fig. 3. The ontology section representing toponyms, along with a running example.

*Epistole, Egloge, and Quaestio de Aqua et Terra*<sup>14</sup>. We chose the works of this author for the fundamental impact that they had on Italian and international literature. Furthermore, the scholars involved in the IMAGO project are authoritative experts in Dante Alighieri's literary production, thus they started to collect the toponyms cited in his works.

Figure 3 shows the ontology section representing the toponyms. A toponym is identified as an instance of the class `Toponym` that we created as a subclass of the class `E41 Appellation`. Each instance has an `rdfs:label` that links the toponym IRI with a string that represents its human-readable name. In our example, Dante Alighieri's toponym we want to represent is the Adriatic Sea, *Adria* in Latin. The class `Place` is linked to the class `Toponym` by the property "is identified by toponym" that we defined as a subproperty of `P1 is identified by`. Each instance of the class `Place` is identified by a Wikidata IRI, and each IRI is linked through the corresponding Pleiades instance using the property `OWL:sameAs`, which indicates that two IRIs actually refer to the same place. In the example, the place linked to *Adria* is identified with the corresponding Wikidata IRI (i.e. <https://www.wikidata.org/wiki/Q13924>), and also to the corresponding Pleiades IRI (i.e. <https://pleiades.stoa.org/places/1004>). For each place, the geographic coordinates are reported. To represent the geographic coordinates, we used the class `E94 Space Primitive`. The class `Place` is linked to the class `E94 Space Primitive` by the property `P168 place is defined by`. The class `F2 Expression`, which represents the work in which the toponym is cited, is linked to the class `Place` by the property `P67 refers to`. In our example, the work in which the Adriatic Sea is cited is the *Egloge*, identified with an Wikidata IRI (i.e. <https://www.wikidata.org/entity/Q921273>). The class `F2 Expression` is linked to the class `Toponym` by the property `P106 is composed of`. To provide the textual context of a toponym along with the textual place in which the toponym is cited, we used the property `R15 has fragment`, which links the classes `F2 Expression` and `E90 Symbolic Object`. Thus the `E90 Symbolic Object` is related to the textual context (a string) through

<sup>14</sup><https://dante.princeton.edu/pdp/>

the property P190 has symbolic content and to the textual place (a string) through the property :has textual place we defined. In our example, the textual place in which *Adria* is cited is "Eg. IV 68" (that means that the toponym is reported in the *Egloge*, 4th piece, versus 68) and the textual content is the sentence: "*Emilida qua terminat Adria terram*". Finally, the class *Toponym* is linked by the property "has preferred identifier" to the *Vocabolario Dantesco Latino (VDL)*<sup>15</sup>, a web resource that reports linguistic explanations of the toponyms cited by Dante Alighieri.

The screenshot shows a web form with the following fields and content:

- Author:** Petrus Paulus Vergerius, n. 23-07-1370, m. 08-07-1444
- Work:** De situ veteris et inclyte urbis Rome
- Toponym:** Rome, Aquileia, ADD TOPONYM
- Genre:** Antiquarian and archaeological treatises, ADD GENRE

Below the form is a blue box containing a list of four manuscript entries, each with a right-pointing arrow:

- Manuscript 1
- Manuscript 2
- Manuscript 3
- Manuscript 4

At the bottom of this box is a button labeled "INSERT SHEET".

Fig. 4. The start page of the IMAGO Web tool.

#### 4 THE ONTOLOGY POPULATION

To populate the ontology, we developed a semi-automatic Web tool to allow scholars to insert knowledge through a user-friendly interface. The tool was created to reduce the time to insert knowledge and to avoid the insertion of mistakes thanks to the use of predefined lists of works, authors, libraries, places, and literary genres. The geographic coordinates of the places are also automatically assigned. The labels and the IRIs associated with the resources included in these predefined lists are extracted from the datasets reported in Section 3. Figure 4 shows the main interface of the tool. Through this interface, the scholars can insert the title of a work, the author, the literary genre and the toponyms which are cited in the work. Figure 5 shows the interface to add data about a manuscript. A similar interface was developed for the print editions. Each field of these interfaces maps a class of the IMAGO ontology.

Once the data about a work is inserted through the tool, this is encoded as an OWL graph and stored in a triple store. The overall process is described in Figure 6. The data collected by the tool is exported as a JSON object. Indeed our software uses a JSON schema to represent the data [4], structured according to the IMAGO ontology classes. The JSON object is processed by a Java software, which transforms it into an OWL graph encoded in RDF/XML and Turtle formats. This software carries out its task by relying on the Apache Jena library<sup>16</sup>. The graph is finally stored into a Fuseki triple

<sup>15</sup><https://www.vocabolariodantescolatino.it/>

<sup>16</sup><https://jena.apache.org/>

store<sup>17</sup>, and it can be queried through a SPARQL endpoint<sup>18</sup>.

Manuscript 1

<b>Author</b>	
<b>Work</b>	Eiusdem P.P. Vergeri Iustinopolitani De situ veteris et inclytae urbis Romae
<b>Place of Library</b>	Parigi
<b>Library</b>	Bibliothèque nationale de France
<b>Signature</b>	Lat. 5879
<b>Folios</b>	7r-10r
<b>Incipit dedication / proem</b>	
<b>Explicit dedication / proem</b>	
<b>Incipit of the text</b>	Dici solet et habet certam res ipsa ratione
<b>Explicit of the text</b>	Civitas quoque Leonina tres habet portas de quarum nominibus
<b>Date</b>	sec. XVI
<b>Decoration / Iconographic apparatus</b>	Il capolettera dell'opera presenta una decorazione a bianchi girari; sono assenti, relativamente al De situ veteris, annotazioni marginali come glosse e notabilia. L'opera è trascritta da un'unica mano.
<b>Manuscript link</b>	<a href="https://gallica.bnf.fr/ark:/12148/btv1b9068286n/f14.item">https://gallica.bnf.fr/ark:/12148/btv1b9068286n/f14.item</a>
<b>Description of manuscript link</b>	<a href="https://archivesetmanuscrits.bnf.fr/ark:/12148/cc64871c">https://archivesetmanuscrits.bnf.fr/ark:/12148/cc64871c</a>
<b>Secondary sources</b>	Smith 1926,
<b>Notes</b>	Il manoscritto, miscellaneo, tramanda anche le seguenti opere: Petri Pauli Vergeri Justinopolitani, Liber de politia Venetorum eorumque legibus ac moribus, urbis situ, naturaque regionis (ff. 2r-6v); Anonymi chronica Italica (ff. 12r-30v); Historiae ecclesiasticae omnium temporum libri sex (ff. 31r-147v); Sexi Rufini, Viri consularis historiae romanae breviarum (ff. 148r-155r); Benvenuti de Rambaldis Imolensis, Liber augustalis de vitis Caesarum (ff. 155r-170r).
<b>Annotator</b>	*****
<b>Last change</b>	14/05/2022

DELETE EDIT

Fig. 5. Example of manuscript annotation form of the IMAGO Web tool.

At the current stage of the project, our KB includes 250 works, 206 authors and 614 libraries, and the scholars have started to insert detailed knowledge about manuscripts and printed editions of these works. The KB also includes ten different literary genres, four types of editions, and six ecdotic typologies. For what concerns the toponyms, at the moment, we only collected those cited by Dante Alighieri in his Latin works, which are 88 corresponding to 80 places. In our case study, we used DanteSearch<sup>19</sup> to automatically retrieve the linguistic data related to the toponyms. DanteSearch is a Web resource that allows querying a complete corpus of Dante Alighieri's works. The corpus is lemmatized and grammatically and syntactically annotated. Using an API<sup>20</sup> we developed to query DanteSearch, for each toponym, we were able to automatically retrieve the textual context in which the toponym occurs and the number of occurrences. Once this data was retrieved, we automatically inserted it in our KB. Table 3 summarises the data about the toponyms of our case study.

<sup>17</sup><https://jena.apache.org/documentation/fuseki2/>

<sup>18</sup><https://imagoarchive.it/sparql/>

<sup>19</sup><https://dantesearch.dantenetwork.it/>

<sup>20</sup><http://vdl.isti.cnr.it/dantesearch.html>

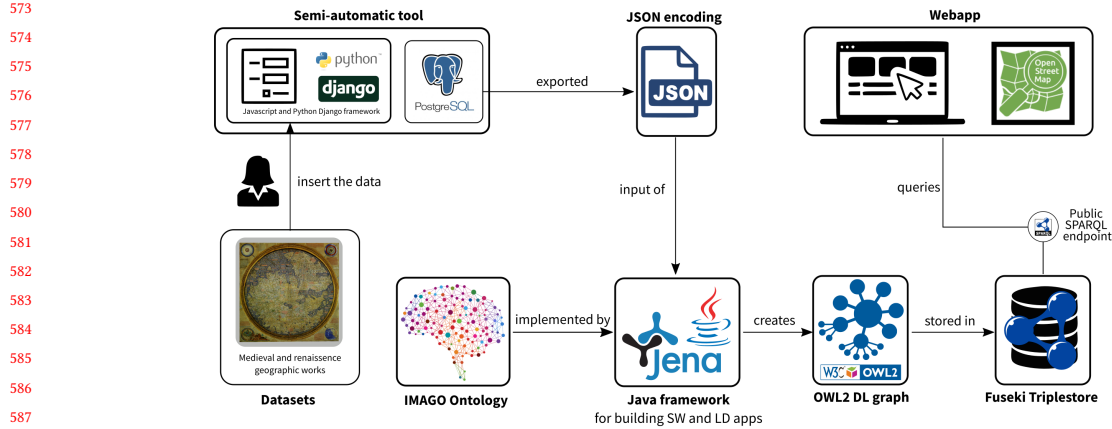


Fig. 6. The architectural schema of the IMAGO project.

Table 3. Data about Dante Alighieri’s toponyms collected in the IMAGO KB.

Data	Total Number
Toponyms	88
Places	80
Works	5
Textual contexts	136

The population tool is open source [31], as well as the Java software that creates the OWL graph [30]. The tool and the Java software are based on the ontological model we made freely available online<sup>21</sup>. Hence, researchers can adapt them to other domains of interest or extend them by adding additional pieces of knowledge. Once the knowledge is collected, and the graph is created, the data is published online as LOD. Having the IMAGO dataset available as LOD allows enriching the knowledge collected in the IMAGO KB by linking it to other datasets. At the same time, the LOD dataset can be used and exploited within other research projects. Indeed, the corpus of geographical knowledge collected in the IMAGO project is entirely new in the Digital Humanities domain. Furthermore, until now, the knowledge about medieval geographical works was dispersed on paper books and unavailable in digital format. Finally, we used Pellet, an open source OWL DL reasoner [34], to check the consistency of the IMAGO ontology to make sure to retrieve consistent knowledge as the result of the SPARQL queries. Indeed, by performing semantic queries on the Fuseki SPARQL endpoint, we are able to discover new knowledge starting from the knowledge explicitly represented in the data. For example, we can retrieve all works reported in the same manuscript, the manuscripts collected in the same library, the centuries in which the production of the manuscripts was more relevant, the manuscripts that describe a specific place or how many times Dante Alighieri cited a place.

## 5 KNOWLEDGE VISUALISATION OF THE DATA RELATED TO THE TOPONYMS

The final aim of the IMAGO project is to create a web application that allows visualising and exploring the data collected in the KB. This knowledge should be accessed by scholars but also by students and general users. For this reason, we

<sup>21</sup><https://imagoarchive.it/documentation/doc/index-en.html>

implemented some user-friendly visualisations to access the data. The design of the visualisation components was defined after the analysis of the requirements we collected through several interviews with the scholars involved in the project. The first type of visualisation we implemented is the presentation of the toponyms on a map<sup>22</sup>. To implement this map, we used the Leaflet library, an open-source JavaScript library for mobile-friendly interactive maps<sup>23</sup>. This library was already successfully used within the Mapping the Enlightenment project for visualising, exploring, and analysing the intellectual and geographical networks developed by Greek-speaking scholars of the Ottoman Empire during the 17th and 18th centuries [33]. Leaflet allows the creation of geographic maps consisting of several layers. In our case, the first layer is the map provided by OpenStreetMap. The second layer shows the pins that identify the toponyms. In particular, Leaflet takes as input the JSON file produced by our tool. The JSON contains all the information about the toponyms. Each toponym is reported as a pin on a map associated with a place that is uniquely identified by its geographic coordinates. Figure 7 shows the map of the toponyms reported by Dante Alighieri in his Latin works.

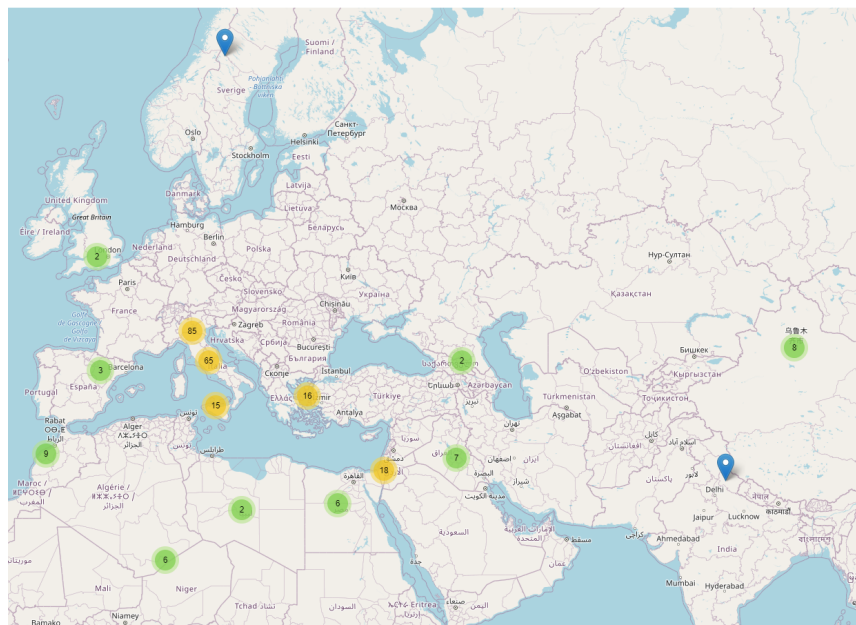


Fig. 7. The toponyms reported by Dante Alighieri in his Latin works.

Clicking on each toponym, a pop-up appears reporting the following data:

- the corresponding linguistic form;
- the work containing the toponym;
- the total number of toponym occurrences in the work;
- the textual place;
- the textual context;
- the link to the corresponding Wikipedia page;
- the link to the corresponding Pleiades page.

<sup>22</sup><https://imagoarchive.it/en/toponyms/>

<sup>23</sup><https://leafletjs.com/>

677 Figure 8 shows the pop-up corresponding to *Adria* containing all the data reported above.

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

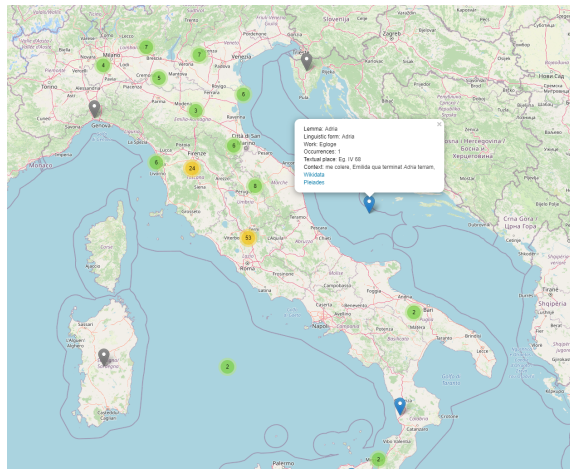


Fig. 8. The data associated to *Adria* reported in a pop-up.

698

699 A toponym can be cited in more than one work or more than once in the same work. This is the case of the toponym  
700 *Florence*, which is cited in different works but also more than once in the same work (e.g. *De Vulgari Eloquentia*,  
701 *Monarchia*). The map reported several pins on the Florence place, one for each citation of a toponym in Dante's works  
702 (Figure 9-a). Clicking on each pin, a pop-up reporting the data associated with the selected toponym citation is shown  
703 (Figure 9-b).

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

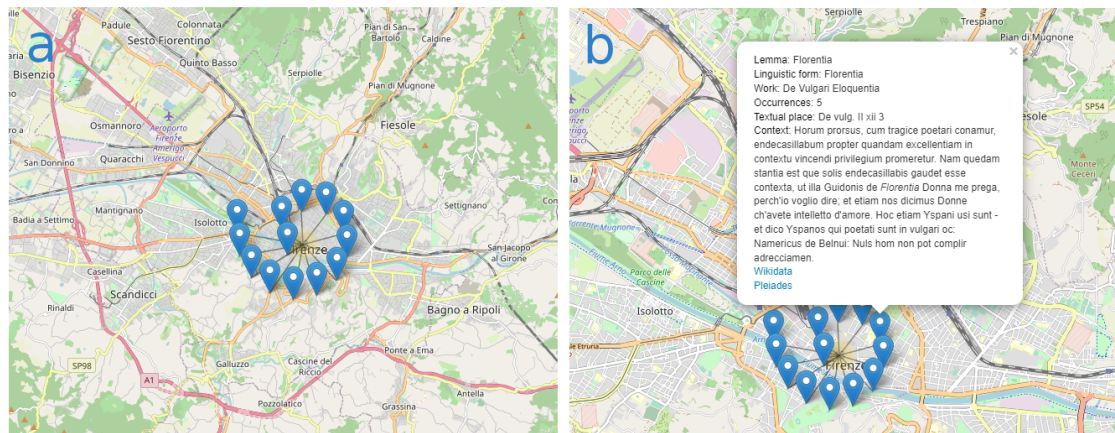


Fig. 9. a. Multiple pins associated to Florence; b. The pop-up reporting the data associated with a citation of the toponym Florence in the *De Vulgari Eloquentia*.

In addition to the map, we designed and implemented an interface that allows users to access the collected data through predefined SPARQL queries. The SPARQL queries are able to retrieve the following data:

Manuscript submitted to ACM

- 729 (1) Selecting a toponym in a list<sup>24</sup>, the user can retrieve:
- 730 • the works in which it appears, the textual contexts, the textual places;
  - 731 • the works in which it appears and the number of occurrences in those works;
  - 732 • the related place along with the geographic coordinates and the corresponding links to Wikidata and Pleiades
  - 733 pages;
  - 734 • the authors who cite the toponym and the works in which it appears.
- 735 (2) Selecting a work in a list<sup>25</sup>, the user can retrieve:
- 736 • the cited toponyms and the number of citations;
  - 737 • the cited places and the corresponding toponyms, which can be more than one, e.g. the toponyms *Ausonia* and
  - 738 *Hesperia* for Italy.
- 739 (3) In the case that the KB includes more than one author, selecting an author in a list, the user can retrieve:
- 740 • all the toponyms cited by the author;
  - 741 • all the works and for each of them the list of the cited toponyms;
  - 742 • for each work, the total number of the cited toponyms.

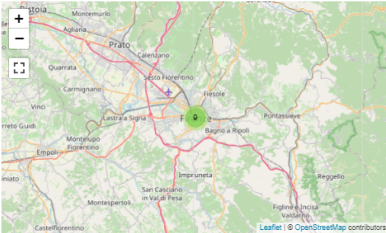
743 The results of all these queries are shown to the user in the form of maps, tables and CSV files.

744 **Search by toponym**

745 Search for **textual contexts**, **textual places** and **works** in which a toponym appears.

746 Select a toponym from the list above

747 Florentia



748 Place associated with the toponym

English	Florence
Italian	Firenze
Latin	Florentia
Coordinates	POINT(11.254166666 43.771388888)
Wikidata place	<a href="https://www.wikidata.org/entity/Q2044">https://www.wikidata.org/entity/Q2044</a>
Pleiades place	<a href="https://pleiades.stoa.org/places/413138">https://pleiades.stoa.org/places/413138</a>

749 Occurrences by work

Author	Work	Occurrences
Dante Alighieri	De Vulgari Eloquentia	4
Dante Alighieri	Epistole	5

750 Download as CSV

751 Works, textual places and contexts

752 Fig. 10. The "Search by Toponym" interface showing some results associated to the toponym *Florentia*

753 <sup>24</sup><https://imagoarchive.it/en/toponyms/toponyms.html>

754 <sup>25</sup><https://imagoarchive.it/en/toponyms/works.html>

781 **Search toponyms by work**  
 782 Select a work and visualize the corresponding **toponyms**, **occurrences**, **textual places**, and **textual contexts**.  
 783 Select a work from the list above  
 784

785 Toponyms and their occurrences

786 Toponym	787 Occurrences
788 Gades	1
789 Ganges	1
Mantua	1
Verona	1

790 Download as CSV

791 Toponyms, textual places and contexts

792 Toponym	793 Wikidata place	794 Textual place	795 Context
796 Mantua	<a href="https://www.wikidata.org/entity/Q6247">https://www.wikidata.org/entity/Q6247</a>	797 Questio 2	798 Manifestum sit omnibus vobis quod, existente me <i>Mantua</i> , questio quedam exorta est, que dilatata multobens ad apparentiam magis quam ad veritatem, indeterminata restabat.
799 Gades	<a href="https://www.wikidata.org/entity/Q15662">https://www.wikidata.org/entity/Q15662</a>	800 Questio 54	801 Nam, ut comunitur ab omnibus habetur, hec habitabilis extenditur per lineam longitudinis a <i>Gadibus</i> , que supra terminos occidentales ab Hercule positos pontur, usque ad hostia fluminis <i>Ganges</i> , ut scribit Orosius. Que quidem longitudo tanta est, ut occidente sole in equinoctiali existente illis qui sunt in altero terminorum, ortur illis qui sunt in altero, sicut per eclipsim lune compertum est ab astrologis. Igitur oportet terminos predicte longitudinis distare per cxxx gradus, que est dimidia distantia totius circumferentie.
802 Ganges	<a href="https://www.wikidata.org/entity/Q65089">https://www.wikidata.org/entity/Q65089</a>	803 Questio 54	804 Nam, ut comunitur ab omnibus habetur, hec habitabilis extenditur per lineam longitudinis a <i>Gadibus</i> , que supra terminos occidentales ab Hercule positos pontur, usque ad hostia fluminis <i>Ganges</i> , ut scribit Orosius. Que quidem longitudo tanta est, ut occidente sole in equinoctiali existente illis qui sunt in altero terminorum, ortur illis qui sunt in altero, sicut per eclipsim lune compertum est ab astrologis. Igitur oportet terminos predicte longitudinis distare per cxxx gradus, que est dimidia distantia totius circumferentie.
805 Verona	<a href="https://www.wikidata.org/entity/Q2028">https://www.wikidata.org/entity/Q2028</a>	806 Questio 87	807 Determinata est hec phylsophia dominante invicto domino, domino Cane Grandi de Scala pro imperio sacrosancto Romano, per me Dantem Alagherium, phylsophorum minimum, in incilla urbe <i>Verona</i> , in sacello Helene gloriose, coram universo clero Veronensi, preter quosdam qui nimis caritate ardentis, aliorum rogamina non admittunt, et per humilitatis virtutem Spiritus Sancti pauperes, ne aliorum excellentiam probare videntur, sermonibus eorum interesse refugunt.

808 Download as CSV

809 Fig. 11. The "Search toponym by work" interface showing the results associated to the work *Questio de Aqua et Terra*

810 Figure 10 shows the interface from which a user can select a toponym and visualise the information related to the place associated with the toponym and the occurrences of the toponym in Dante's works. The Figure reports some pieces of knowledge related to the toponym *Florentia*. Figure 11 shows the interface from which, by selecting one of Dante's works, a user can retrieve information about the toponyms that occur in that work, the number of occurrences, the textual places and contexts. In particular, Figure 11 reports the information related to the toponyms that occur in the *Questio de Aqua et Terra*.

## 814 6 CONCLUSION AND FUTURE WORK

816 In this paper, we have presented how we have used the Semantic Web technologies for representing the knowledge about the toponyms cited in the medieval Latin geographical works collected within the IMAGO project. To reach this aim, we have developed an CIDOC-CRM-based ontology that allows representing the knowledge about this literature. In particular, the ontology allows representing different pieces of knowledge about toponyms, e.g. the place which the toponym refers to, the work in which it is cited, the textual context in which it appears, and the total number of citations in a work. Each resource collected in the IMAGO is represented by an IRI. In order to maximise the interoperability of our KB, we retrieve and reuse IRIs from existing and *standard* KBs, i.e. Wikidata, Pleiades, Mapping Manuscript Migrations. The ontology is expressed in OWL 2 DL and we have designed and developed a Web tool to populate it. On top of the ontology, different visualisations were created, which allow scholars, students and general users to retrieve and access the knowledge about the toponyms. In particular, we have implemented a visualisation of the toponyms on a map, where clicking on the pin that identified the toponym, a pop-up appears that shows all the information about the toponym we collected. Another visualisation is in the form of tables and CSV files that aggregate data about a toponym (e.g. works in which it appears, number of occurrences, the corresponding place). As case study, we considered the Manuscript submitted to ACM



833 toponyms cited by the Italian poet Dante Alighieri in his Latin works. However, one of the final aims of IMAGO project  
834 is to create a complete index of the toponyms cited in the medieval and renaissance Latin geographical works. The  
835 main advantages of using a Semantic Web approach instead of a traditional database to create such an index are the  
836 following: (i) the ontology allows to easily add classes and properties, thereby refining the ontology, (ii) the ontology  
837 allows representing a huge variety of semantic relationships easily, (iii) the ontology can be linked to other ontologies  
838 to extend the represented domain, (vi) any user can download and use the knowledge base freely, using the paradigm of  
839 Linked Open Data. Eventually, the digitisation of the knowledge about medieval toponyms collected in the IMAGO KB  
840 and their visualisation in the form of maps, tables and CSV files allow the scholars to have a complete overview of this  
841 data. For example, it is possible to easily carry out a first analysis of the spatial distribution of the toponyms of each  
842 author, visualising them on a map. Furthermore, scholars can study the evolution of geographical knowledge over time,  
843 automatically analysing which and how many toponyms were cited by different authors who lived in different times or  
844 by the same author in his/her works written in different years. For example, the visualisations of the results of the case  
845 study reported in this paper, are able to provide an overview of Dante Alighieri's geographical knowledge and allow  
846 understanding of how this knowledge evolved in his works. We think that it is a precious information for scholars, but  
847 also for students and generally for users interested in Dante's literature. To facilitate the perception of the evolution of  
848 geographical knowledge over time, as future work, we have also planned to show the temporal knowledge related to  
849 the ancient manuscripts and digital editions on a timeline. In [2] we have reported the results of a first evaluation of  
850 the ontological model we performed. As future work, we will carry out a complete assessment of the ontology, which  
851 include not only the model but also the KB of the project. In particular, we plan to conduct two different types of  
852 assessment: an automatic assessment and an assessment involving users. For the first type of assessment, we plan to use  
853 the automatic OntoQA system [36] that evaluates both the model and the KB. For the second type of assessment, we  
854 plan to propose a specific questionnaire to the scholars who are currently populating the ontology. After the analysis of  
855 the assessment results, if necessary, we will review and extend our ontology.

856 For what concerns the toponyms, we have planned to expand our KB with those cited in all works we are collecting.  
857 Furthermore, we are working to implement a data visualisation in the form of network graphs that will allow users to  
858 identify the toponyms shared by different works. The long term aim of the project is to develop a Web application that  
859 allows retrieving and consulting all the data collected in the IMAGO KB in a user-friendly way.

## 860 REFERENCES

- 861 [1] MG Aune. 2011. The Itineraries of William Wey. *Renaissance Quarterly* 64, 1 (2011), 310.
- 862 [2] Valentina Bartalesi, Daniele Metilli, Nicolò Pratelli, and Paolo Pontari. 2022. Towards a knowledge base of medieval and renaissance geographical  
863 Latin works: The IMAGO ontology. *Digital Scholarship in the Humanities* 37, 1 (2022), 34–50. <https://doi.org/10.1093/lc/fqab060>
- 864 [3] Valentina Bartalesi and Nicolò Pratelli. 2021. The IMAGO Project: Towards a Knowledge Base of Medieval and Renaissance Geographical Works  
865 (extended abstract). In *Proceedings of the International Joint Workshop on Semantic Web and Ontology Design for Cultural Heritage co-located with the  
866 Bolzano Summer of Knowledge 2021 (BOSK 2021), Virtual Event / ozen-Bolzano, Italy, September 20-21, 2021 (CEUR Workshop Proceedings, Vol. 2949)*,  
867 Antonis Bikakis, Roberta Ferrario, Stéphane Jean, Béatrice Markhoff, Alessandro Mosca, and Marianna Nicolosi Asmundo (Eds.). CEUR-WS.org,  
868 <http://ceur-ws.org/Vol-2949/short1.pdf>
- 869 [4] Valentina Bartalesi, Nicolò Pratelli, and Emanuele Lenzi. 2022. *JSON schema of the internal IMAGO tool data representation*. Retrieved December 19,  
870 2022 from [https://imagoarchive.it/imago\\_json.html](https://imagoarchive.it/imago_json.html)
- 871 [5] Sebastian Barzaghi, Monica Palmirani, and Silvio Peroni. 2020. Development of an ontology for modelling medieval manuscripts: The case of  
872 Progetto IRNERIO. *Umanistica Digitale* 9 (2020), 117–140. <https://doi.org/10.6092/issn.2532-8816/11187>
- 873 [6] Florian Bauer and Martin Kaltenböck. 2011. Linked open data: The essentials. *Edition mono/monochrom, Vienna* 710 (2011).
- 874 [7] Christian Bizer, Tom Heath, and Tim Berners-Lee. 2011. Linked data: The story so far. In *Semantic services, interoperability and web applications:  
875 emerging concepts*. IGI global, 205–227.
- 876 [8] Nathalie Bouloux. 1999. *Culture et savoirs géographiques dans l'Italie du XIVe siècle*. Ph. D. Dissertation. Paris 1.

- 885 [9] Toby Burrows, Douglas Emery, Mitch Fraas, Eero HyvÄänen, Esko Ikkala, Mikko Koho, David Lewis, Andrew Morrison, Kevin Page, Lynn Ransom,  
886 Emma Thomson, Jouni Tuominen, Athanasios Velios, and Hanno Wijsman. 2020. Mapping Manuscript Migrations: Digging into Data for Researching  
887 the History and Provenance of Medieval and Renaissance Manuscripts: White Paper.
- 888 [10] Daniel K Connolly. 1999. Imagined pilgrimage in the itinerary maps of Matthew Paris. *The Art Bulletin* 81, 4 (1999), 598–622. <https://doi.org/10.2307/3051336>
- 889 [11] Domenico Defilippis. 2009. Da Flavio Biondo a Leandro Alberti. Corografia e antiquaria tra Quattro e Cinquecento, Atti del Convegno di Studi  
890 (Foggia, 2 febbraio 2006), a cura di D. Defilippis. (2009).
- 891 [12] Linfang Ding, Guohui Xiao, Albulen Pano, Claus Stadler, and Diego Calvanese. 2021. Towards the next generation of the LinkedGeoData project  
892 using virtual knowledge graphs. *Journal of Web Semantics* 71 (2021), 100662. <https://doi.org/10.1016/j.websem.2021.100662>
- 893 [13] Martin Doerr. 2003. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. *AI magazine* 24, 3  
894 (2003), 75–75. <https://doi.org/10.1609/aimag.v24i3.1720>
- 895 [14] Martin Doerr, Chryssoula Bekiari, Patrick LeBoeuf, and Bibliothèque nationale de France. 2008. FRBRoo, a conceptual model for performing arts. In  
896 *Proceedings of the 2008 Annual Conference of CIDOC, Athens*. 15–18.
- 897 [15] Fredo Erxleben, Michael Günther, Markus Krötzsch, Julian Mendez, and Denny Vrandečić. 2014. Introducing Wikidata to the linked data web. In  
898 *Proceedings of the International semantic web conference*. Springer, 50–65. [https://doi.org/10.1007/978-3-319-11964-9\\_4](https://doi.org/10.1007/978-3-319-11964-9_4)
- 899 [16] Stefanie Gehrke, Eduard Frunzeanu, Pauline Charbonnier, and Marie Muffat. 2015. Biblissima Prototype on Medieval Manuscript Illuminations and  
900 their Context.. In *Proceedings of the International Workshop Semantic Web for Scientific Heritage at ESWC Conference*. 43–48.
- 901 [17] Sergio Gensini. 2000. *Viaggiare nel Medioevo*. Pacini Centro di Studi sulla civiltà del tardo Medioevo, Pisa San Miniato.
- 902 [18] Pascal Hitzler, Markus Krotzsch, and Sebastian Rudolph. 2009. *Foundations of semantic web technologies*. Chapman and Hall/CRC.
- 903 [19] Eero Hyvönen, Esko Ikkala, Mikko Koho, Jouni Tuominen, Toby Burrows, Lynn Ransom, and Hanno Wijsman. 2021. Mapping Manuscript Migrations  
904 on the Semantic Web: A Semantic Portal and Linked Open Data Service for Premodern Manuscript Research. In *Proceedings of the International  
905 Semantic Web Conference*. Springer, 615–630. [https://doi.org/10.1007/978-3-030-88361-4\\_36](https://doi.org/10.1007/978-3-030-88361-4_36)
- 906 [20] Anna Jordanous, K Faith Lawrence, Mark Hedges, and Charlotte Tupman. 2012. Exploring manuscripts: Sharing ancient wisdoms across the semantic  
907 web. In *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics*. 1–12. <https://doi.org/10.1145/2254129.2254184>
- 908 [21] Tobias Kohr, Kai-Christian Bruhn, Ashish Karmacharya, Christophe Cruz, and Frank Boochs. 2013. An SDI for archaeological data with a RESTful  
909 interface to semantically modeled information. In *Geoinformatik*.
- 910 [22] Paul Oskar Kristeller. 1963. *Iter Italicum: Alia itinera 1. Australia to Germany*. Vol. 3. Warburg Institute.
- 911 [23] Markus Krötzsch. 2012. OWL 2 profiles: An introduction to lightweight ontology languages. In *Reasoning Web International Summer School*. Springer,  
912 112–183. [https://doi.org/10.1007/978-3-642-33158-9\\_4](https://doi.org/10.1007/978-3-642-33158-9_4)
- 913 [24] Deborah L McGuinness, Frank Van Harmelen, et al. 2004. OWL web ontology language overview. *W3C recommendation* 10, 10 (2004), 2004.
- 914 [25] Enrico Menestò. 1993. Relazioni di viaggi e di ambasciatori. *Cavallo, Guglielmo; Leonardi, Claudio; Menestò, Enrico (dir. da), Lo spazio letterario del  
915 Medioevo* 1 (1993), 535–600.
- 916 [26] Friar Odoric. 2002. The Travels of Friar Odoric. *Edited by Henry Yule and Paolo Chiesa. Translated* (2002).
- 917 [27] Jeff Z Pan. 2009. Resource description framework. In *Handbook on ontologies*. Springer, 71–90. [https://doi.org/10.1007/978-3-540-92673-3\\_3](https://doi.org/10.1007/978-3-540-92673-3_3)
- 918 [28] Paolo Pontari. 2016. «Nedum mille qui effluerunt annorum gesta sciamus». L'Italia di Biondo e l'invenzione del Medioevo. (2016).
- 919 [29] August Potthast. 1970. *Repertorium fontium historiae medii aevi*. (1970).
- 920 [30] Nicolò Pratelli and Valentina Bartalesi. 2022. *The Java software to create the IMAGO OWL graph*. Retrieved December 19, 2022 from <https://github.com/prate91/imagotriplifier>
- 921 [31] Nicolò Pratelli, Valentina Bartalesi, and Emanuele Lenzi. 2022. *IMAGO Annotation Tool*. Retrieved December 19, 2022 from [https://github.com/prate91/IMAGO\\_annotation\\_tool](https://github.com/prate91/IMAGO_annotation_tool)
- 922 [32] Pat Riva and Maja Žumer. 2017. FRBRoo, the IFLA Library Reference Model, and now LRMoo: a circle of development. (2017).
- 923 [33] Vasileios Routsis, Irene Goudarouli, and Manolis Patiniotis. 2017. Mapping the Enlightenment: Intellectual Networks and the Making of Knowledge  
924 in the European Periphery. In *DH*.
- 925 [34] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. 2007. Pellet: A practical owl-dl reasoner. *Journal of Web  
926 Semantics* 5, 2 (2007), 51–53. <https://doi.org/10.1016/j.websem.2007.03.004>
- 927 [35] Claus Stadler, Jens Lehmann, Konrad Höffner, and Sören Auer. 2012. Linkedgeodata: A core for a web of spatial open data. *Semantic Web* 3, 4 (2012),  
928 333–354. <https://doi.org/10.3233/SW-2011-0052>
- 929 [36] Samir Tartir, I Budak Arpinar, Michael Moore, Amit P Sheth, and Boanerges Aleman-Meza. 2005. OntoQA: Metric-based ontology quality analysis.  
930 (2005).
- 931 [37] Sabine Tittel. 2022. Towards an Ontology for Toponyms in Nepalese Historical Documents. In *Proceedings of the Workshop on Resources and  
932 Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia within the 13th Language Resources and Evaluation Conference*.  
933 7–16.
- 934 [38] Denny Vrandečić. 2013. The rise of Wikidata. *IEEE Intelligent Systems* 28, 4 (2013), 90–95. <https://doi.org/10.1109/MIS.2013.119>
- 935 [39] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten,  
936 Luiz Bonino da Silva Santos, Philip E Bourne, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific  
937 data* 3, 1 (2016), 1–9. <https://doi.org/10.1038/sdata.2016.18>