An Ontology for the Iconographic Sources of *La Divina Commedia*⋆

Andrea De Domenico[1], Domenico Cantone[2], and Marianna Nicolosi-Asmundo[2]

[1] Scuola Superiore di Catania, University of Catania, Italy
email: andreadedomenico@studium.unict.it
[2] Department of Mathematics and Computer Science, University of Catania, Italy
email: {domenico.cantone,marianna.nicolosiasmundo}@unict.it

Abstract. In recent years, scholars have been very interested in the relationship between Dante Alighieri and the figurative arts, leading the way to new research paths that have enriched this topic and made it much more complex and interesting. Currently, this knowledge is disseminated on several books and research papers, making it unmanageable to systematically overview the literary and artistic background of Dante and to gain a sound understanding of how this background was gradually constructed over time. We present OntoComedySources, an ontology for the primary sources of Dante Alighieri's *Commedia*, both literary and iconographic, with a focus on the latter. We developed our ontology using the Web Ontology Language 2 (OWL 2), according also to several foundational ontologies widely used in the ambit of the digital humanities, such as the CIDOC Conceptual Reference Model (CRM) and the Semantic Publishing and Referencing Ontologies (SPAR Ontologies). Our main goal is to support academics in developing and consulting a digital encyclopedia of *La Divina Commedia*'s primary sources. In addition, we expect that OntoComedySources can easily be adapted and extended to other works and other authors.

1 Introduction

The relationship between *La Divina Commedia* and figurative arts is one of the research topics on which scholars have been very active in recent years. Unfortunately, along with information held in the more acknowledged literary sources, such a relationship remains mainly implicit to contemporary readers. It is the task of specialized scholars to analyze and reveal these hidden connections.

Semantic web offers well established tools and technologies to deal with such issues. Indeed, it provides formal languages for knowledge representation, such as RDF Schema and the Web Ontology Language, which enable one to connect data and information from different sources at a global level, thus enhancing coherence, integration, and dissemination of knowledge. Moreover, dedicated automated reasoning systems permit to verify the consistency of representations,

query datasets, and infer implicit information from what has been already defined. Formal representations of application domains are called *ontologies*.

Recently, ontologies have been used to represent and organize primary sources in Dante Alighieri's works [3, 2]. Such contributions consider five principal works of Dante, namely *Convivio*, *De vulgari eloquentia*, *De Monarchia*, *Rime*, and *Vita Nova*, focusing on their literary sources.

In this note, we present OntoComedySources,[3] an ontology aimed at representing the iconographic and literary sources of Dante Alighieri's *Commedia*, with particular emphasis on the former ones. Our ontology categorizes the various kinds of sources and connects them with the corresponding text fragments in the *Commedia*. To this purposes, we make use of different standard foundational ontologies in the digital humanities such as the CIDOC Conceptual Reference Model (CRM), the Open Annotation Core Data Model (OA), the Dublin Core ontology (DC), and some modules from the Semantic Publishing and Referencing (SPAR) Ontologies.

We chose CIDOC to model knowledge related to iconographic works, also in consideration of the fact that it was successfully used in the definition of ontologies devoted to iconography [7, 9]. Concerning the representation of information related to literary sources, we used OA, DC, and SPAR inspired by the analysis carried out by [2].

The primary goal of our efforts is to provide the means necessary to support academics in developing and consulting a digital encyclopedia of *La Divina Commedia*'s primary sources. Digital representation of information sources can be a powerful instrument in the understanding of classical literature. In the case of the *Commedia*, it would provide significant benefits in the diffusion, use, and comprehension of knowledge hidden in several publications, archives, and commentaries fragmented across the globe and the web.

With little effort, our ontology could be readapted and extended to other literary works and other classical authors.

2 Preliminaries

2.1 Semantic web and ontologies

Semantic web, conceived by Tim Berners-Lee and presented in 2001, is an evolution of the web in which machine-readable data can be queried and manipulated by software agents on behalf of human agents. In such a vision, web information is explicitly interconnected and equipped with meaning, so that it can be automatically processed by machines, and data can be accessed, integrated, and modified at a global level, resulting in increased coherence and dissemination of knowledge. In addition, by means of automated reasoning procedures, it is possible to extract and process implicit information present in data, thus permitting to gain a deeper knowledge of the domain.

[3] https://github.com/AndreaDeDomenico97/OntoComedySources

The *Word Wide Web Consortium* (W3C) recommends OWL, a family of knowledge representation languages based on *Description Logics* (DLs) [1], as a standard tool for the semantic web, specifically for the representation of ontologies. An ontology is a formal description of a domain of interest carried out by combining three basic syntactic categories: *entities*, *expressions*, and *axioms*, forming the logical part of ontologies, namely what ontologies can express and which types of inferences can be drawn [12, 11]. Ontologies can also be combined together in order to describe more complex domains. A specific type of ontologies, namely *upper-level* or *foundation ontologies* [12], is designed to model high-level and domain-independent categories about the real world. This provides general terms which are used to connect domain-specific ontologies (lower-level ontologies) allowing one to reach a broader semantic interoperability. Some examples of foundational ontologies are given in Section 2.2.

OWL, currently in its version 2.1, is grounded on *triples* or *statements*, each one representing an atomic unit. These are ways to connect either two entities, or an entity and a data-value, each one represented by an *Internationalized Resource Identifier* (IRI), i.e., a sequence of characters that unambiguously identifies a resource within a specific context.

Entities represent the primitive terms of an ontology and are identified in a unique way. They are *individuals* (actors), *properties* (actions), and *classes* (sets of actors with common features). Properties are of two types: object-properties and datatype-properties. Object-properties relate pairs of individuals, whereas datatype-properties relate individuals with some data type values, such as strings and numbers.

OWL expressivity can be extended by adding Horn-like rules by means of the *Semantic Web Rule Language* (SWRL).[4] Rules have the form of an implication between an antecedent (body) and consequent (head), intending that whenever the conditions specified in the antecedent hold, the conditions specified in the consequent must hold as well.

2.2 Foundational ontologies used in OntoComedySources

In order to implement a good modeling approach and facilitate future integrations with other ontologies, we constructed OntoComedySources according to the following foundational ontologies, widely used in the digital humanities.

- The *CIDOC Conceptual Reference Model* (CRM) is the international standard for the controlled exchange of cultural heritage information since 2006. It provides a general specification which can be adopted in any cultural heritage context to construct a semantic web-based information system, to serve as a guide for good practices of conceptual modelling, and to improve information sharing. There are several institutions that successfully implement CIDOC such as galleries, libraries, museums, archives, as well as any other cultural environment based on cultural heritage data that publishes

[4] https://www.w3.org/Submission/SWRL/

and shares its information in the semantic web formats. The CIDOC core covers several general aspects of cultural information, such as material and immaterial entities, events, space, and time. Such general concepts can effectively represent iconographic works.

- The *Dublin Core* (DC) ontology is part of the DCMI project (Dublin Core Metadata Initiative), now an ISO standard. It is composed by a core of terms essential to the description of any digital material, as well as physical resources such as artworks or books. We used it to define properties for designating the creator of a literary or iconographic work and the author of a commentary.
- The *Semantic Publishing and Referencing* (SPAR) is a family of ontologies for the description of any aspect of bibliographic resources. Specifically, we used the following SPAR ontologies:
  - FRBR (*Functional Requirements for Bibliographic Records model*): a model developed for the description of essential aspects of bibliographic records;
  - FaBiO (*FRBR-aligned Bibliographic Ontology*): an ontology describing entities that are published or potentially publishable (e.g., journal articles, conference papers, books) and that contain or are referred to by bibliographic references;
  - C40 (*Citation Counting and Context Characterisation Ontology*): an ontology that permits to record in-text citations of a cited source, together with their textual citation contexts;
  - DoCO (*Document Components Ontology*): an ontology that provides a structured vocabulary written of document components, both structural (e.g., paragraph, section, chapter) and rhetorical (e.g., introduction, discussion, acknowledgements);
  - eFRBRoo: an OWL-DL implementation of an object oriented version of FRBR called FRBRoo.
- The *Open Annotation Core Data Model* (OA) specifies a framework for modeling associations between related resources or annotations. It is helpful for representing commentaries to literary works and for explicating relationships between them and the resources they refer to.

3 The ontology OntoComedySources

As illustrated in Fig. 1, the structure of the *Commedia* is modeled by the classes *Verse*, *Canto*, and *Cantica*. These are subclasses of *doco:Line*, *doco:Chapter*, and *fabio:Book*, respectively. Naturally enough, a verse is contained in a canto and, in its turn, a canto is a part of a cantica. This is expressed by the object properties *isVerseOf*, *isCantoOf*, and *inCanticum*, along with their inverses shown in Fig. 1. The properties just mentioned are all subproperties of the two FRBR properties *frbr:hasPart* and *frbr:isPartOf*.

In the following diagrams, classes are represented with oval-shaped borders, while primitive data-types are delimited by rectangular boxes. Object properties

and data properties are denoted with solid lines, whereas dotted lines designate subclass relationships.

The name of a class or property drawn from an existing foundational ontology is written in the form $\langle prefix \rangle : \langle name \rangle$, where the prefix, which is omitted when referring to new entities, stands for the namespace of the ontology.

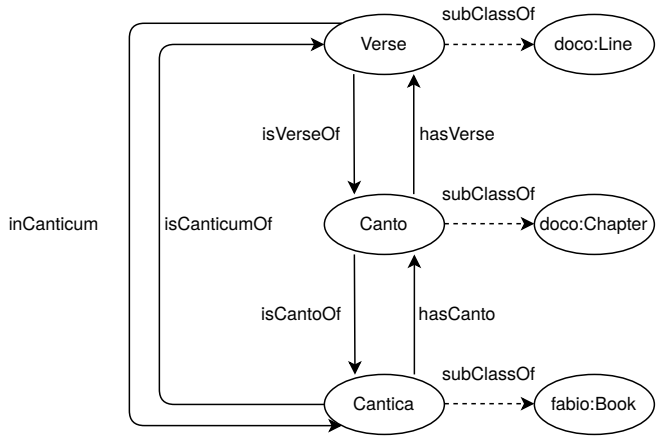

Fig. 1. Structure of *La Divina Commedia*.

The class *Fragment*, subclass of *efrbroo:ExpressionFragment*, occupies a central role in the ontology, as each of its instances refers to a specific textual fragment of *La Divina Commedia*. Its content is expressed by the data property *bodyChars*.

Based on the class *oa:Selector* from the Open Annotation Core Data Model, we are able to make explicit the exact position of the corresponding fragment in the text (see Fig. 2). Specifically, if a certain fragment is located in the interval from the i-th to the j-th character, we can use the properties *oa:start* and *oa:end* to model it. Since the specific type of selector depends on the format of the information which one is working with, we used the subclass *oa:TextPositionSelector*.

Every fragment has a starting verse and an ending verse; furthermore, in principle, a fragment could refer to smaller fragments within it. Such knowledge is represented by means of the properties *startVerse*, *endVerse*, and *composedOf*. Notice that both the first two properties are subproperties of *oa:hasSource*.

To explain the connection between a text fragment and the literary or iconographic source it refers to, we used the class *oa:Annotation*. As shown in Fig. 3, every instance of the class *oa:Annotation* is paired with an instance of the class *oa:TextualBody* via the property *oa:hasBody*. A textual body has a date and a textual content. Every annotation is written by a specialized scholar and in our ontology we specify this fact with the property *dc:contributor*, whose range is the class *Scholar*, subclass of *foaf:Person*. The classes *oa:Annotation* and

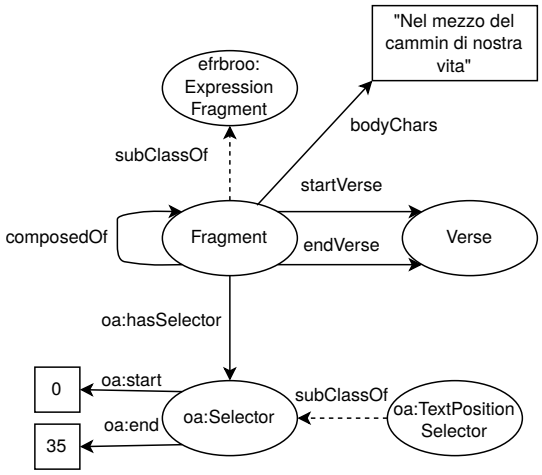

Fig. 2. Representation of text fragments.

oa:TextualBody are related to the class *Fragment* via the properties *oa:hasTarget* and *hasCitingFragment*, respectively.

The association between a fragment and the work cited in it is modeled by the property *c4o:cites*, belonging to the *Citation Counting and Context Characterisation Ontology* (C4O). The range of the property *c4o:cites* is the class *efrbroo:Work*. In its turn, such a class is paired with the class *Author* through the property *dc:creator* and with the class *skos:Concept* via the property *dc:subject*. The instances of the class *efrbroo:Work* are the literary and iconographic sources we are interested in, while the class *skos:Concept* is used to specify the thematic area of the related work (e.g., Patristics). The meaning of the other classes is self-evident.

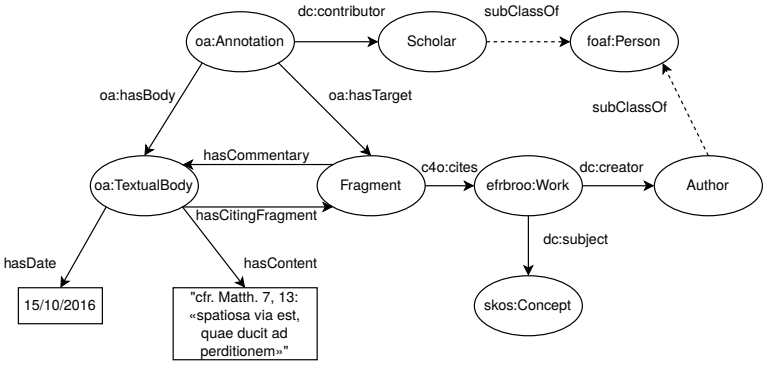

Fig. 3. How fragments, annotations, and cited sources are linked with each other.

As depicted in Fig. 4, the class *efrbroo:Work* is a superclass of both classes *LiteraryWork* and *IconographicWork*. Besides, *LiteraryWork* contains the class *Manuscript*, while *IconographicWork* is a superclass of the four classes *Fresco*, *Mosaic*, *Miniature*, and *PanelPainting*. The properties *citeAsLiterarySource* and *citeAsIconographicSource* are both subproperties of *c4o:cites*. The former has *LiteraryWork* as range, whereas the range of the latter is *IconographicWork*.

To specify the spatial dimensions of an iconographic work, we make use of the class *Length*, subclass of *cidoc:Dimension*, along with the properties *hasDimX* and *hasDimY*. Each instance of *Length* is associated with an instance of the class *cidoc:MeasurementUnit* via the property *cidoc:hasUnit*.

To model the datation of an iconographic work, we utilized the CIDOC class *cidoc:TimeSpan* and the CIDOC property *cidoc:hasTimeSpan*, as well as the properties *startDate* and *endDate*, both subproperties of *cidoc:ongoingThroughout*. To register where the works are currently preserved, we employed the classes *Site* and *Region*, subclasses of *cidoc:Place*. The properties *inSite* and *inRegion*, whose meaning is obvious, are subproperties of *cidoc:hasCurrentPermanentLocation*.

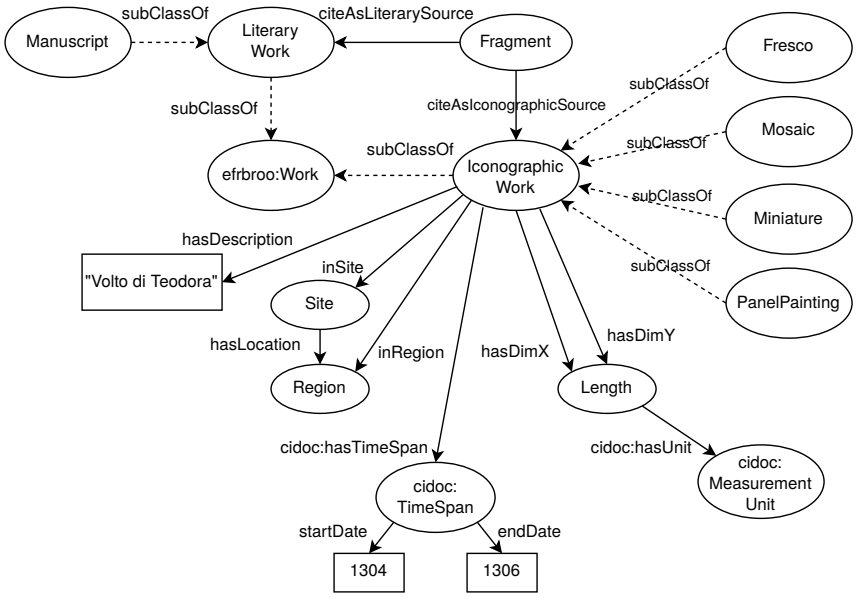

Fig. 4. Focus on the subclass *IconographicWork*.

A miniature is a small illustration contained in an ancient manuscript. In OntoComedySources, this is modeled by the property *isContainedIn*. Since we often have no knowledge concerning the author of a manuscript, we included the class *Antiquary* and the property *acquiredBy* to designate the antiquarian responsible for the discovery and/or the preservation of a given manuscript.

As illustrated in Fig. 5, the ontology provides us with the entities to describe the materials employed in the execution of the mosaics and paintings, along with the techniques used. The properties *typeOfTesserae*, *typeOfSubstrateMosaic* and *typeOfSubstratePainting* connect the classes *Mosaic* and *PanelPainting* with the class *Material*. All of the three properties have *cidoc:consistOf* as a superproperty.

In our ontology, the class *cidoc:DesignOrProcedure* is a superclass of the class *Technique* which, in its turn, has as subclasses the classes *PaintingTechnique* and *MosaicTechnique*; these are linked with the classes *PanelPainting* and *Mosaic* by the properties *usedTechniquePainting* and *usedTechiqueMosaic*, respectively.

In order to differentiate among the different types of mosaics, we exploit the class *MosaicType*, subclass of *cidoc:Type*, and the property *typeOfMosaic*, subproperty of *cidoc:hasType*.

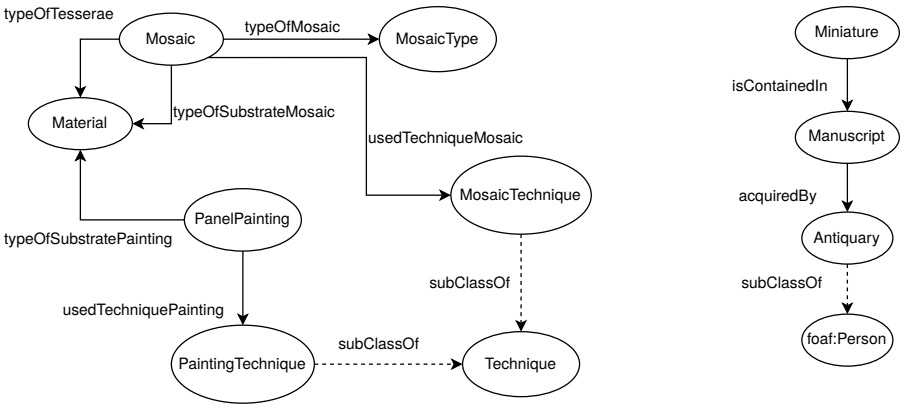

Fig. 5. Properties of *IconographicWork*'s subclasses.

OntoComedySources contains 46 classes, 76 properties, and 209 individuals. It also includes some SWRL rules permitting to draw simple inferences concerning geographical places in relation with iconographic sources, fragments of the *Commedia* with their corresponding commentaries, the position of a verse inside the *Commedia*, and the authors cited in a given canto. Some of the rules are illustrated in Fig. 6. In particular, $hasIconographicSource(a, b)$ means that a canto a contains b as an iconographic source, $hasIconographicSourceCitedIn(a, b)$ holds if a region a contains an iconographic source cited in a canto b, and finally $citedIn(a, b)$ means that a canto b contains a work from an author a as a primary source.

Fig. 7 shows an example of an inference drawn from the third rule in Fig. 6 stating that if canto $?a$ has as iconographic source work $?b$ and work $?b$ is housed in region $?c$, then region $?c$ contains an iconographic source cited in canto $?a$. Instantiating the variables $?a$, $?b$, and $?c$ with the individuals *Pu_9*, standing for

Purgatory's canto IX, *IWorkMo1*, representing the mosaic depicted in Fig. 8, and *Ravenna*, respectively, we infer that the town of Ravenna stows an iconographic source referenced in Purgatory's canto IX.

The ontology has been classified using the DL reasoners Hermit [10] and Pellet [14].

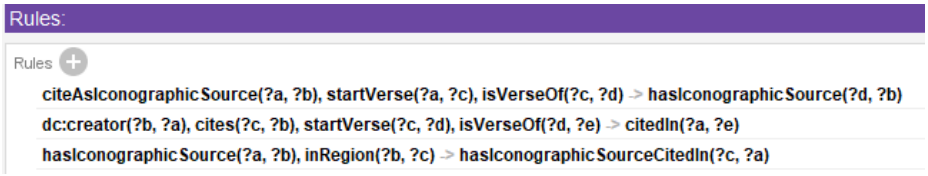

Fig. 6. Some SWRL rules concerning iconographic sources and cited authors.

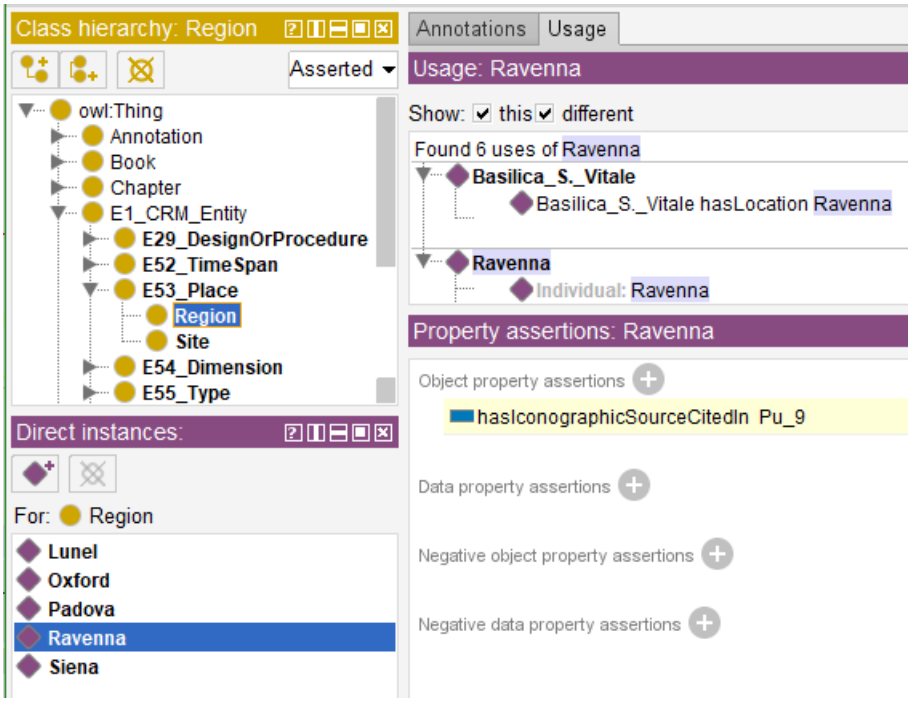

Fig. 7. Example of an inference drawn from the SWRL rules of the ontology.

3.1 Representing an iconographic source

For demonstration purposes, some literary and iconographic sources were added to OntoComedySources. All the literary sources pertain to *Inferno*'s canto V, and they were drawn from two commentaries by the scholars Anna Maria Chiavacci Leonardi [8] e Umberto Bosco [5], respectively.

Iconographic sources are harder to find in mainstream commentaries on the *Commedia*. We extracted our sources from a research paper and a monography written by Alessandro Benucci [4] and Laura Pasquini [13], respectively.

An example of an iconographic source is displayed in Fig. 8. Specifically, Fig. 8 shows a byzantine mosaic panel, executed in 520-547 A.D. and currently conserved in the Basilica of San Vitale, which depicts Empress Theodora. The corresponding fragment is from Purgatory's canto IX. The scholar suggests that, when writing about the aurora, Dante was inspired by the ornaments in Theodora's crown.

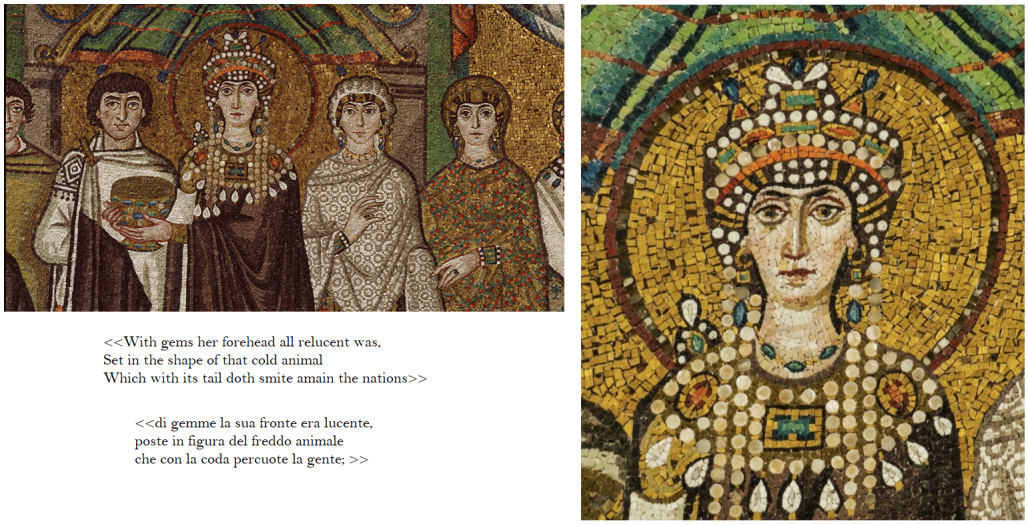

Fig. 8. Example of an iconographic source, taken from [13].

4 Conclusions

In this note we presented OntoComedySources, an OWL 2 ontology for the primary sources of Dante Alighieri's *Commedia*. We modeled OntoComedySources according to the standard CIDOC CRM ontology in order to accurately represent iconographic sources, and we used the foundational ontologies DC, SPAR, and OA to describe literary sources.

It is widely recognized among scholars that Dante's *Commedia* was influenced by the iconographic art of its time. Indeed, in medieval times, figurative arts played a significant role in educating the population, as it was mostly illiterate and did not have access to literary texts. Accordingly, we plan to expand the ontology in order to represent iconographic sources even more broadly, following some recent works on ontologies devoted to iconography [7, 9]. As in Dante's time there were several famous sculptors in Tuscany, and in *La Divina Commedia* there are descriptions of high and low relieves, we also intend to include in our study primary sources concerning sculptural art and, possibly, epigraphy. Concerning the latter issue, we will consider the possibility to integrate OntoComedySources with the ontology EpiONT, presented in [6].

Acknowledgments

We gratefully acknowledge partial support from project "STORAGE—Università degli Studi di Catania, Piano della Ricerca 2020/2022, Linea di intervento 2".

References

1. F. Baader, I. Horrocks, C. Lutz, and U. Sattler. *An Introduction to Description Logic*. Cambridge University Press, 2017.
2. V. Bartalesi and C. Meghini. Using an Ontology for Representing the Knowledge on Literary Texts: the Dante Alighieri Case Study. *Semantic Web Journal, vol. 8, no. 3, pp. 385–394*, 2017.
3. V. Bartalesi, C. Meghini, P. Andriani, and M. Tavoni. Towards a Semantic Network of Dante's Works and Their Contextual Knowledge. *Digit. Scholarsh. Humanit., vol. 30 (suppl. 1), pp. i28–i35*, 2015.
4. A. Benucci. La processione santa di Paradiso XXIII. *Ortodossia ed eterodossia in Dante Alighieri: atti del convegno di Madrid, 5-7 novembre 2012, pp. 361–378*, 2012.
5. U. Bosco and G. Reggio. *La Divina Commedia. Inferno*. Le Monnier, 1981.
6. D. Cantone, S. Cristofaro, M. Nicolosi-Asmundo, F. Prado, D.F. Santamaria, and D. Spampinato. An EpiDoc ontological perspective: the epigraphs of the Castello Ursino Civic Museum of Catania via CIDOC CRM. *Archeologia e calcolatori, vol. 30, 139–157*, 2019.
7. N. Carboni and L. de Luca. An Ontological Approach to the Description of Visual and Iconographical Representations. *Heritage, vol. 2, no. 2, pp. 1191–1210*, 2019.
8. A. M. Chiavacci Leonardi. *La Divina Commedia. Inferno*. Mondadori, 2014.
9. R. Gartner. Towards an ontology-based iconography. *Digital Scholarship in the Humanities*, 2020.
10. B. Glimm, I. Horrocks, B. Motik, G. Stoilos, and Z. Wang. HermiT: An OWL 2 Reasoner. *Journal of Automated Reasoning*, 53(3):245–269, 2014.
11. T. Hofweber. *Logic and Ontology*. The Stanford Encyclopedia of Philosophy (Summer 2018 Edition). Edward N. Zalta (ed.), 2018.
12. D. Oberle, N. Guarino, and S. Staab. *What is an ontology?* in Handbook on Ontologies, 2nd edition, S.Staab and R. Studer eds. Springer, 2009.

13. L. Pasquini. *Iconografie dantesche. Dalla luce del mosaico all'immagine profetica*. Longo Editore Ravenna, 2008.
14. E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, and Y. Katz. Pellet: A practical OWL-DL reasoner. *J. Web Sem.*, 5(2):51–53, 2007.