Modelling Lexicographic Resources using CIDOC-CRM, FRBRoo and Ontolex-Lemon

Fahad Khan[1][0000−0002−1551−7438] and Ana Salgado[2,3][0000−0002−6670−3564]

[1] Istituto Di Linguistica Computazionale 'A. Zampolli', Italy
[2] NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Portugal
[3] Academia das Ciências de Lisboa, Portugal
fahad.khan@ilc.cnr.it, anacastrosalgado@gmail.com

Abstract. The article describes a new approach to the modelling and publication of lexicographic resources, including retro-digitised dictionaries, as linked data. This approach is based on the use of the CIDOC-CRM aligned FRBRoo ontology together with the Ontolex-Lemon vocabulary and its follow up lexicographic module, *lexicog*. After introducing the TEI-based distinction between different views on lexicographic resources, we discuss Ontolex-Lemon and CIDOC-CRM and FRBRoo. Next, we look at some motivating use cases before introducing our approach. Finally, we model one of these use cases in more depth using this approach.

Keywords: data modelling · dictionary · linked data · tei · lexicographic resources

1 Introduction

In this article, we propose an approach to the modelling and publication of retro-digitised dictionaries (and related resources) as linked data based on the use of the CIDOC Conceptual Reference Model (CIDOC-CRM) aligned FRBRoo ontology together with the Ontolex-Lemon vocabulary and its follow-up module, lexicog (all introduced in Section 2). To this end, we propose several new classes and properties to act as a bridge between the first resource and the latter two vocabularies (see Section 4). We motivate our approach in Section 2.1 by examining the different levels of descriptions that lexicographic works usually contain and that this approach is specifically designed to model. Our analysis is based on several representative use cases (presented in Section 3).

2 Background

2.1 Different Views on Lexicographic Resources

This article discusses the modelling of lexicographic resources as linked data. The term lexicographic resources includes so-called *machine-readable dictionaries* [6], that is, lexical resources that are either digital versions of previously

existing paper dictionaries (so-called retro-digitised dictionaries) or updated digital editions of such dictionaries. In addition, we also classify as lexicographic resources born-digital lexical resources which are structured or organised according to the conventions of traditional dictionaries (or which contain significant portions organised as such). In broader terms, we define a lexicographic resource as a language resource a significant portion of which consists of *a set of lexicographic articles or entries* written and organised according to some suitable lexicographic criteria. Prima facie, there are two main approaches which we can take in modelling lexicographic resources in general and retro-digitised dictionaries in particular. We can view a dictionary primarily as a cultural/textual artefact with its own specific publishing history and its own verbal expression and visual arrangement of the linguistic content contained within it. Or we can prioritise the linguistic content instead, ignoring how it is presented and the exact sequence of words used in, for example, the definitions of articles. Of course, there is also a third (potentially more verbose) approach, that is, to do both at the same time and make sure both kinds of information are aligned. The distinction made in Chapter 9 ("Dictionaries") of the Text Encoding Guidelines (TEI) between the *typographical*, *editorial* and *lexical* views of a dictionary is particularly useful in this discussion. These views are defined as follows: the typographical view concerns the layout of individual pages; the editorial view is concerned with the properties of a text modelled as a sequence of tokens; and the lexical view is concerned with the conceptual or linguistic content of a dictionary as a whole, as well as its individual entries[1]. In addition to these views, TEI also offers extensive provision, in the *teiHeader* element, for including metadata about an original resource to be modelled (e.g., who the authors were and when it was published) and about the process of its digitisation as well as the curation of that process.

In this work, we will concentrate on the modelling of lexicographic resources as linked data to be published on the Semantic Web. This means that the distinctions made in the TEI guidelines are not sufficient for our purposes since the data modelling affordances that TEI-XML offers are different from those offered by linked data and the Semantic Web. This is due both to the expressive possibilities offered by knowledge representation formalisms, such as the **Resource Data Framework (RDF)** and the **Web Ontology Language (OWL)**, as well as to the ready availability of tools and relevant ontologies and vocabularies in those formalisms. All of this offers data modellers enhanced possibilities for representing the conceptual relationships between different levels of description in lexicographic resources in a more explicit and, hence, more machine-actionable way. Provision for the modelling of the TEI lexical view in a linked data lexico-

---

[1] With the typographical view we would be interested in e.g., the position of line breaks in a text, or the visual arrangement of entries on any single page; with the editorial view we are interested in such things as which words are used in the description of the article and in which order; finally in the lexical view we're interested in information about the given domain of a specialised lexical unit (a term) or sense of a lexical unit, or that a given headword is a "verb".

graphic resource is already covered in large part by **Ontolex-Lemon** (which is explored in some detail in Section 2.2) with the latter providing a common data model for the creation of linked data lexicons. Such lexicons can potentially contain information derived from various sources (including, but not limited to, lexicographic resources), and so the use of a common model that imposes conceptual restrictions on how lexical data is represented ultimately facilitates the interoperability of such datasets. Such lexicons can potentially contain information derived from various different sources (including but not limited to lexicographic resources) and so the use of a common model that imposes conceptual restrictions on how lexical data is represented[2] ultimately facilitates the interoperability of such datasets. However, this presents us with the problem of translating the surface representation of this same information (this representation corresponding to the typographic and editorial views mentioned above) into linked data as well as of relating these different levels of description together with each other and with the lexical view – something that is not always straightforward[3]. We are often also interested in numerous other details related to the conception and publication of dictionaries. This is true of historically important dictionaries and/or dictionaries that have gone through significant changes across their different editions.

Broadly speaking then, we are talking here about the distinction between the *linguistic content* of a dictionary text and its *metadata* (we would argue that the latter category also includes the TEI typographic and editorial views). Our approach (which as far as we are aware is fairly novel with regard to lexicographic resources) is to model this complexity using the top-level ontology CIDOC-CRM and the CIDOC-aligned ontology FRBRoo along with Ontolex-Lemon and *lexicog*. In other words, we first seek to model separately and then to integrate together the various different aspects of a lexicographic resource using a conceptual modelling approach that uses the ontologies/vocabularies mentioned above along with various others available on the Semantic Web. We propose to do this in a way that makes the semantics of the data as explicit and therefore as machine-actionable as possible. To motivate our approach (described in more detail in Section 4) and to help the reader better appreciate some of the challenges involved, we will look at one specific kind of use case in detail (in Section 3). Before that, however, we will provide brief introductions to some of the models and vocabularies just mentioned.

---

[2] For instance a headword, corresponding to a lexical entry in Ontolex-Lemon, can only be associated with a single part of speech

[3] At the same time however our task is simplified by the fact that the information in a dictionary (as represented by the lexical view) often (though not always) tends to be of a standard form and is organised in a fairly standardised way. At least this is the case to an extent which isn't true of for example biographies or scientific papers, or most other kinds of text where we might be interested in annotating both structural aspects as well as the truth values of the claims being made in the text.

2.2 Ontolex-Lemon and Lexicog

The *Lexicon Model for Ontologies* of the **W3C Ontology-Lexicon community group**, also known as Ontolex-Lemon, is a linked data native model for the creation of lexical resources[4] [9]. Ontolex-Lemon was originally developed to facilitate the addition of linguistic information to Semantic Web ontologies. However, it quickly became a de facto standard for the modelling of lexicons in RDF with or without the ontological component as originally envisaged. Inspired by previous lexical standards such as the Lexical Markup Framework [6], the vocabulary contains classes such as **Lexicon**, **Lexical Entry**, **Lexical Sense**, **Form** and **Word** (the names of the classes are fairly self-explanatory).The popularity of the Ontolex-Lemon vocabulary as well as the need to deal with cases where there is a substantial difference between the organisation of entries on a page and the modelling proposed by Ontolex-Lemon[5], soon led to calls for an extension of the original model specifically tailored to the representation of lexicographic resources. The *OntoLex Lemon Lexicography Module* (lexicog) was subsequently developed in response [3]. Figure 1 shows the main classes and properties of the new module. The class **Lexicographic Resource**, as distinct from the **Lexicon** class in Ontolex-Lemon (included in the *lime* sub-module of the latter), consists of a collection of **Entry** individuals where each such individual is "a structural element that represents a lexicographic article or record as it is arranged in a source lexicographic resource"[6]. The latter is also a subclass of the class **Lexicographic Component** whose individuals are defined "structural element[s]" representing "the (sub-)structures of lexicographic articles providing information about entries, senses or sub-entries". Individuals of the class **Entry** then are linked to Ontolex-Lemon entries via the **entry** object property with the latter relating **Entry** to "an element that represents the actual information provided by that component in the lexicographic resource". The *lexicog* module makes a good start at addressing the modelling issues raised by the presence of different dictionary views as identified by the TEI guidelines. However as we will argue in Section 3 there are important kinds of information relevant to modelling lexicographic resources which are still not covered by the combination of Ontolex-Lemon and *lexicog*.

2.3 CIDOC-CRM and FRBRoo

The **CIDOC Conceptual Reference Model** (CIDOC-CRM)[7] is an event-based conceptual model or ontology that was initially developed for the cultural heritage domain with the aim of facilitating the harmonisation and integration

---

[4] The commnunity page is here: https://www.w3.org/community/ontolex/; the guidelines can be found here: https://www.w3.org/2016/05/ontolex/.

[5] One well-known example of this is when dictionaries violate the Ontolex-Lemon model constraint requiring one part of speech per **Lexical Entry**

[6] This and the other following *lexicog* quotes are taken from the guidelines to be found at https://www.w3.org/2019/09/lexicog/
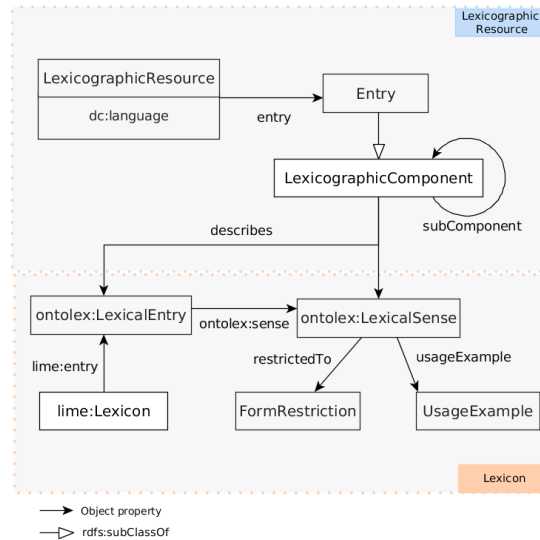
[7] http://www.cidoc-crm.org/

Fig. 1. the OntoLex Lemon Lexicography Module

of data arising from different cultural heritage organisations [5]. CIDOC-CRM makes a clear distinction between *endurants* and *perdurants*, i.e., between entities with a persistent identity that may orginate, survive or terminate in events, and temporal concepts which *happen* rather than maintaining a stable identity over time. The **Functional Requirements for Bibliographic Records** (FRBR) entity relationship model, instead, was developed by the **International Federation of Library Associations and Institutions** as a conceptual model for the classification of intellectual products in bibliographic databases and library catalogues [4]. FRBR is well known for its four-fold classification of bibliographic entities based on the distinction between a *Work*, *an Expression*, *a Manifestation* and *an Item*. The distinction is an extremely pertinent one for our discussion as we will see in Section 3. We refer in what follows to the version of this distinction introduced in **FRBRoo**, an ontology based on FRBR and which is harmonised with CIDOC-CRM [2][8]:

– **F1 Work**: "[C]omprises distinct concepts or combinations of concepts identified in artistic and intellectual expressions [...] The substance of Work is ideas" [2]. Note that in the case of dictionaries this would encompass the TEI lexical view;

– **F2 Expression**: "[C]omprises the intellectual or artistic realisations of works in the form of identifiable immaterial objects, such as texts, poems [...] or any combination of such forms that have objectively recognisable structures"[2].

---

[8] Note that FRBRoo class identifiers are prefixed by the letter 'F' and a number.

In the case of dictionaries we claim that this description encompasses the TEI editorial view;

– Originally one class in the FRBR model, **Manifestation**, this latter corresponds to two separate classes in FRBRoo: **F3 Manifestation Product Type** and **F4 Manifestation Singleton**. The former class is defined as "[defining] all of the features or traits that instances of F5 Item normally display in order that they may be recognised as copies of a particular publication" [2]; the latter as "[comprising] physical objects that each carry an instance of F2 Expression, and that were produced as unique objects" [2]. In the case of dictionaries **F3 Manifestation Product Type** encompasses the TEI typographic view;

– **F5 Item**: "[C]omprises physical objects" [2] such as specific physical copies of dictionaries kept at libraries or academic institutions. This class is associated with the kind of metadata information that is usually contained within the TEI header element.

These FRBRoo classes (together with various others included in the ontology) are ideal for describing those aspects of a lexicographic resource which we mentioned above and which are related to its metadata (including those described in the typographical and lexical views). For instance, we can classify a multi-edition dictionary as a **F15 Complex Work** with its different individual editions classed as instances of **F15 Individual Work**. Each of these different editions can then be described at the level of **F2 Expression** in order to specify, for example, the wording of individual entries. Moreover we can also describe the dictionary at the level of **F3 Manifestation Product Type** in order to specify the content and placement of images and their relation to the text (this is important in the case of illustrated dictionaries). FRBRoo also allows for the modelling of dictionaries which have been translated from one language to another (such as, for instance, the well-known *Liddell-Scott-Jones* Ancient Greek-English lexicon which has been translated into several languages, including Modern Greek). Other than these benefits however, the use of an ontology to model a lexicographic resource, and especially one that has been aligned with a popular top level ontology (CIDOC-CRM in this case), helps to improve the interoperability and re-usability of the resulting dataset.

Once more it is worth emphasising that dictionaries are both *cultural artefacts* and highly structured descriptions of *linguistic phenomena*. It is also worth underlining the importance of being able to relate together these aspects of the same resource. It would be useful, for instance, to be able to represent how lexical content changes over time and across different editions of a dictionary. We should also be able to describe the linguistic claims that a dictionary makes regardless of their truth or falsity (this latter as determined by an editor or curator). This would allow for a higher level description of such claims as well as for the integration of data from different lexical and lexicographic sources and/or the enrichment of single lexicographic resources. In order to further develop this point, and to look at these issues in more detail, we will look at a specific kind of lexicographic data in the next section, namely lexicographic ci-

tations/attestations; this will also provide us with a use case which we will model below using our new approach.

3   Citations and Attestations

Dictionaries very often contain citations under individual entries. This is especially typical of scholarly or encyclopedic dictionaries such as the *Oxford English Dictionary* or the *Liddell-Scott-Jones Ancient Greek-English Lexicon (LSJ)* where individual lexicographic articles are often associated with large numbers of them. Here citations are usually meant to show that a lexical unit, a sense, a form, or indeed any kind of linguistic phenomena is *attested to* in a given text or corpus. Of course these purported citations can sometimes be erroneous. This is the case, for example, where the text cited doesn't actually attest to the phenomenon described due to a misreading of the source or even where it doesn't exist in the form which the lexicographer hypothesised. In these kinds of situation we are concerned with facts that pertain to (at least) two different levels of description: facts at the level of the work (a dictionary citing a text as attesting a given phenomenon) and at the level of content (the phenomenon is or is not attested to by the text in question).

   A couple of examples may be useful here; both of them are taken from [7][9]. In that article we looked at one of the citations in the LSJ lexicographic article for the word ἀνώμᾰλος (*anomalos*). This citation was marked by the abbreviation 'cj'[10] and referred to a critical reconstruction of a work based on existing fragments. The citation was subsequently excluded from another Ancient Greek dictionary which was closely based on the LSJ thus hinting at its unreliability. We can therefore describe this citation both according to which dictionaries or lexical resources it is found in (and which editions) as well as whether the cited text really attests to the phenomenon in question. The other example concerns the Italian verb **riprovare** which means both[11] 'to try something again'[12] and 'to scold, rebuke'. The well known Italian language dictionary *Il vocabolario Treccani* [10] (which we will refer to as Treccani from hereon in) lists these as two separate lexicographic articles: **riprovare**$_1$ ('to try again') and **riprovare**$_2$ ('to scold')[13] The entry for *riprovare*$_1$ cites the motto of the short lived but highly influential 16th scientific society **L'Accademia del Cimento (AdC)**, i.e., ***provando e***

---

[9] Note that although we presented these same examples in that article our approach was different in that work and didn't involve the use of CIDOC-CRM or any upper ontology for structuring information at different levels of description.

[10] This is usually found in critical apparatuses and stands for the Latin *conicit* 'conjectures'.

[11] In fact we are dealing with two homonymous lexical units here.

[12] Deriving in this instance from *provare* which has the meaning of 'to try' and the prefix *ri-* which adds the sense of repetition.

[13] See the online version of both entries can be found here:https://www.treccani.it/vocabolario/riprovare1/, https://www.treccani.it/vocabolario/riprovare2/

riprovando ('trying and trying again'). This motto derives from a terzina of Dante's Divine Comedy (Par. III, 1-3): "Quel sol che pria d'amor mi scaldò 'l petto,/di bella verità m'avea scoverto,/provando e riprovando, il dolce aspetto". The Treccani entry however directs the dictionary user to the second sense of the entry for the verb *provare*[14] 'to try'. There it is explained that although the motto derives from Par. III, 1-3, it does not, in fact, mean the same thing as it does in the latter[15]. The Treccani entry for $riprovare_2$ also cites the use of the gerund form *riprovando* in the same passage of Dante, but in this case it is as an actual attestation. Another authoritative Italian language dictionary, *Il Grande Dizionario della Lingua Italiana* (GDLL) [1][16], cites both the motto of L'Accademia del Cimento and the same terzina from the Divine Comedy in its entry for $riprovare_1$ (and note that $riprovare_1$ and $riprovare_2$ in the GDLL are synonymous with $riprovare_1$ and $riprovare_2$ respectively in Treccani). The GDLL makes the contradictory claim, however, that both cited texts (the AdC motto and the terzina from the Paradiso) *do attest* to the entry in question, namely $riprovare_1$.

We can summarise the claims made above in the following series of statements: **(1)** Treccani's entry for $riprovare_1$ cites the AdC motto; **(2)** Treccani's entry for $riprovare_2$ cites Par. III, 1-3 as an attestation; **(3)** GDLL's entry for $riprovare_1$ cites Par. III, 1-3 and the AdC mottos as attestations; **(4)** $riprovare_2$ is attested by the AdC motto and Par. III, 1-3; **(5)** $riprovare_1$ is attested by the AdC motto but not Par. III, 1-3.

The first three items in the list are true statements about the lexicons which they refer to; they describe the existence of three successful citational ('performative') speech acts. These statements only deal indirectly with lexical units or their usages. Rather, they are primarily concerned with documents or works and the rhetorical/organizational structure pertaining to them. The other two statements, those which we have numbered 4 and 5, instead, describe the direct relationship between an item in a lexicon and a text which evidences it, or better, attests to its past use. The fourth statement is false but the fifth one is true. However in neither case does this follow from the truth (or falsity) of the first three statements. Statements 4 and 5 are at the level of content whereas the first three are at the level of metadata. If we were to create a linked data lexicographic resource combining the two dictionaries listed above (based on a lexical view) and including relevant metadata information about the original sources, then we could include 1,2,3, and 5 as RDF statements but adding 5 would be to add a false statement. In case we are not sure about which claims are true (or which claims we would endorse by creating a curated version of the information in these dictionaries) then we should stay at the metadata level. We could also

---

[14] https://www.treccani.it/vocabolario/provare/

[15] "[M]otto assunto verso il 1666 dall'Accademia fiorentina del Cimento, tratto dalla Divina Commedia (Par. III, 3, dove però significa 'approvando e disapprovando')"

[16] The dictionary is consultable here as a collection of scans: http://www.gdli.it/; the image of the page for riprovare can be found here: http://www.gdli.it/JPG/GDLI16/00727.jpg.

publish the GDLL with provenance metadata of course without curating it, or publish separate statements about single entries or senses as nanopublications [8] but it would also be useful to have the possibility of putting together a single dataset in a way that integrates together this information deriving from multiple sources. In the next section, we look at how we can model this situation with linked data ontologies.

4   Our Proposed Approach

Our intention is to create bridges between the models and vocabularies which we have mentioned so far. We do this through the definition of a number of new classes and properties. This should facilitate a more extensive modelling of lexicographic resources and at multiple levels of description. The first of these new classes is **Lexicographic Work**. This is a subclass of *both* the FRBRoo class **F1 Work** and the Ontolex-Lemon class **Lexicon**. We define it as follows: **Lexicographic Work** *comprises concepts or combinations of concepts for representing/describing the lexicon for a given language community or communities or domain.* As **F1 Work** is a subclass of the CIDOC-CRM class **E89 Propositional Object**[17] we can view individuals of **Lexicographic Work** as sets of propositions about lexemes and related linguistic concepts belonging to a lexicon. The second new class that we introduce is **Lexicographic Expression**. This is a subclass of the FRBRoo class **F2 Expression** and the Ontolex-Lemon class **Lexicographic Resource**. We define it as follows: **Lexicographic Expression** *comprises an intellectual realisation of the description of a lexicon as a structured text.* In other words, it is a text viewed apart from a specific typographic realisation: a sequence of words that has an additional organisation in terms of entries, senses (defined as a sub-part of a lexicographical article that discusses a meaning of a lexical unit), forms, etc. By making these newly defined classes subclasses of both Ontolex-Lemon and FRBRoo classes we can take advantage of properties and related classes from both ontologies (as well as from the CIDOC-CRM). Moreover, and analogously to the *lexicog* class **Entry**, we propose the creation of new subclasses of the *lexicog* class **Lexicographic Component**, corresponding to senses and forms as parts of a typographical view. We also envision an extensive use of other Semantic Web ontologies for provenance metadata and in particular of Prov-O [18] and the **Citation Typology Ontology** (CiTO) (part of the **SPAR** suite of ontologies for the publishing domain[19]). We use the object property **cites** from the latter ontology in our example below[20]. This property has a series of sub-properties corresponding to different kinds of citations (e.g., **cites**

---

[17] Defined as "sets of propositions about real or mental things and that are documented as single units or serve as topic of discourse"[2].

[18] https://www.w3.org/TR/prov-o/

[19] http://www.sparontologies.net/

[20] This property applies to cases where "[t]he citing entity cites the cited entity, either directly and explicitly (as in the reference list of a journal article), indirectly (e.g. by citing a more recent paper by the same group on the same topic), or im-

as source document, **cites for information**). We have decided to define another subproperty of **cites** or more precisely of the CiTO property **cites as evidence**. This new property is **cites as attestation** and covers cases where the citing entity cites the cited entity *as attesting to the existence of a linguistic phenomenon or phenomena*. We also make use of the new *Frequency Attestation and Corpus Information* (FrAC) module being drafted by the W3C OntoLex group in our example encoding. Although this is still currently under development, the module guidelines are in a fairly advanced state of preparation and we have chosen classes and properties in our example encoding whose definitions are, by now, stable[21]. These include **Attestation**, representing attestations of linguistic "observables" in texts and the properties **attestation**, which relates together these observables and the texts which attest them, and **locus** which represents locations in a text. In the rest of this section we describe our proposal for encoding the *riprovare* example from Section 3 using the vocabularies/ontologies we have mentioned above and the new properties we have just defined. In Figure 2 we show the two entries both at the level of the **Lexicographic Work** and at the level of **Lexicographic Expression**. We can use OWL the **sameAs** property to
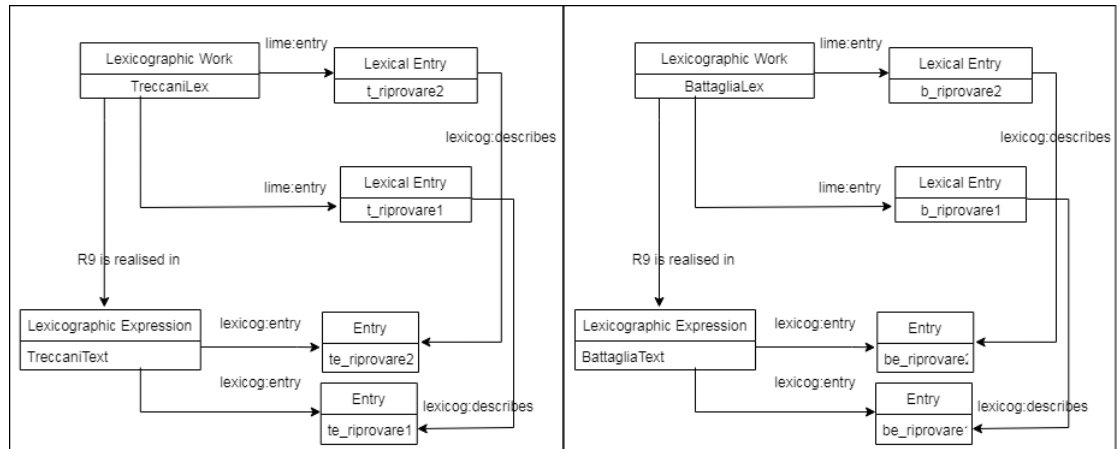

Fig. 2. Dictionaries

link together **t_riprovare1** and **b_riprovare2** (note that we would not do this in the case of the two **Entry** individuals **te_riprovare1** and **be_riprovare2**). In Fig 3 instead we present the relationship between the two entries and the same textual fragment of Dante's Divine Comedy (which we have defined as an individual of the FRBRoo class **Expression Fragment**) as defined above, using the FRaC class

---

plicitly (e.g. as in artistic quotations or parodies, or in cases of plagiarism)". See https://sparontologies.github.io/cito/current/cito.html

[21] The latest draft can be found here: https://github.com/ontolex/frequency-attestation-corpus-information

Attestation and the properties **attestation** and **locus**. Note that we have only represented the fact that GDLL cites the Paradiso as an attestation at the level of Expression and not at the level of Work/Lexicon (although we have done so in the case of Treccani because of the veridicity of the description of the latter).
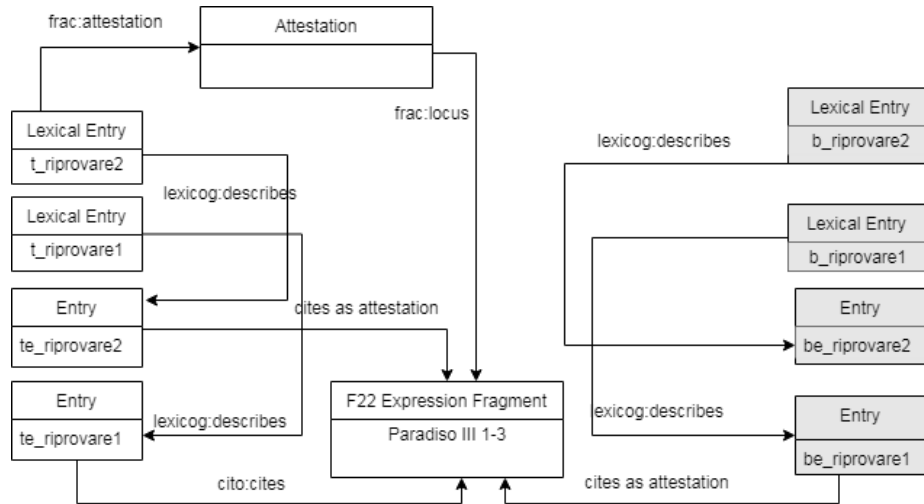

Fig. 3. Attestations

5 Further Work

The work which we have described in this article represents an attempt to use ontologies to model Semantic Web lexicographic resources at different levels of description. We have found CIDOC-CRM and its aligned/related ontologies (together with pre-existing lexical ontologies such as Ontolex-Lemon) to be especially useful in this regard. Indeed we are currently working on representing linguistic hypotheses (such as etymologies) in lexical resources using CRMSci. Furthermore, the authors of the current work are involved in the digitization of three editions of the historically significant Portuguese dictionary, *Diccionario da Lingua Portugueza* (Dictionary of the Portuguese Language), by António de Morais Silva, as part of a Portuguese national project, MOR*Digital*. The intention is to make these dictionaries made available in both TEI-XML as well as in linked data. This project will give us the opportunity to apply and to test the approach which we have proposed in the current article on a large scale, real-life dataset.

Acknowledgements

References

1. Battaglia, S.: Il Grande dizionario della lingua italiana di Salvatore Battaglia. UTET, Torino (1961)
2. Bekiari, C., Doerr, M., Le Boeuf, P., Riva, P., Aalberg, T., Barthélémy, J., Boutard, G., Görz, G., Iorizzo, D., Jacob, M., et al.: FRBR object-oriented definition and mapping from FRBRER, FRAD and FRSAD (Version 3.0) (Sep 2017), http://www.cidoc-crm.org/frbroo/sites/default/files/FRBRoo_V3.0.pdf
3. Bosque-Gil, J., Gracia, J., Montiel-Ponsoda, E.: Towards a Module for Lexicography in OntoLex. In: Proceedings of the LDK 2017 Workshops: 1st Workshop on the OntoLex Model (OntoLex-2017), Shared Task on Translation Inference Across Dictionaries & Challenges for Wordnets. Galway, Ireland (2017), http://ceur-ws.org/Vol-1899/
4. Coyle, K.: FRBR, before and after: a look at our bibliographic models. ALA Editions, an imprint of the American Library Association, Chicago (2016)
5. Doerr, M.: The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. AI Magazine **24**(3), 75–75 (Sep 2003)
6. Francopoulo, G. (ed.): LMF lexical markup framework. Computer engineering and IT series, ISTE Ltd ; John Wiley & Sons, London : Hoboken, NJ (2013), oCLC: ocn812570948
7. Khan, F., Boschetti, F.: Towards a Representation of Citations in Linked Data Lexical Resources. In: Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts (2018)
8. Kuhn, T., Banda, J.M., Willighagen, E., Ehrhart, F., Evelo, C., Malas, T.B., Dumontier, M., Merono-Penuela, A., Malic, A., Poelen, J.H., et al.: Nanopublications: A growing resource of provenance-centric scientific linked data. In: 2018 IEEE 14th International Conference on e-Science (e-Science). p. 83–92. IEEE (2018)
9. McCrae, J.P., Bosque-Gil, J., Gracia, J., Buitelaar, P., Cimiano, P.: The Ontolex-Lemon model: development and applications. In: Proceedings of eLex 2017 conference. pp. 19–21 (2017)
10. Simone, R. (ed.): Enciclopedia dell'italiano: Il vocabolario Treccani. Ist. della Enciclopedia Italiana, Roma, ed. speciale per la libreria edn. (2010), oCLC: 0759152511